LDPC Codes for Partial-Erasure Channels in Multi-Level Memories

Rami Cohen and Yuval Cassuto
Department of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200000, Israel
Email: rc@tx.technion.ac.il, ycassuto@ee.technion.ac.il

***Abstract*—In this paper, we develop a new channel model, which we name the q-ary partial erasure channel (QPEC). QPEC has a q-ary input, and its output is either one symbol or a set of M possible values. This channel mimics situations when current/voltage levels in measurement channels are only partially known, due to high read rates or imperfect current/voltage sensing. Our investigation is concentrated on the performance of low-density parity-check (LDPC) codes when used over this channel, due to their low decoding complexity with iterative-decoding algorithms. We give the density evolution equations of this channel, and develop its decoding-threshold analysis. Part of the analysis shows that finding the exact decoding threshold efficiently lies upon a solution to an open problem in additive combinatorics. For this part, we give bounds and approximations.**

I. Introduction

The advent of non-volatile memories (NVMs) with many levels per cell holds a great promise for increased storage capacity. At the same time, it proves extremely challenging to write and read many-level cells at both high precision and high speeds. As a result, coding is employed to improve the tradeoff between data reliability and access speed (see e.g. [1]). Natural candidates to improve the reliability of non-volatile memories are low-density parity-check (LDPC) codes [2], which offer low complexity of implementation and good performance under iterative decoding [3]. Recent work on the employment of LDPC codes in NVMs, such as [4], focused on the additive white Gaussian noise (AWGN) channel.

In addition to the AWGN and other classical channels, NVMs motivate coding for a diversity of new channels with rich features. Our work here is motivated by a class of channels we call *measurement channels*, which encompass a variety of equivocations introduced to the information by an imperfect read process. This imperfection of the read process comes from either physical limitations or speed constraints. In particular, the channel model we study here – the *q-ary partial erasure channel* (QPEC) – comes from a read process that occasionally fails to read the information at its entirety, and provides as decoder inputs q-ary symbols that are *partially* erased.

Theoretically speaking, the QPEC is an extension of the q-ary erasure channel (QEC) [5], where instead of erasing a full channel symbol, the channel returns a set of $M \le q$ symbols that contains the correct stored symbol and $M-1$ other symbols. Our results on the QPEC include calculating its capacity in Section II, a message-passing decoder in Section III, and analysis and approximation models for its density-evolution formulation in Sections IV and V.

II. QPEC: Q-ary Partial Erasure Channel

A. Channel model

The *Q-ary Partial Erasure Channel* (QPEC) is defined as follows. Let X be the transmitted symbol taken from the alphabet $\mathcal{X} = \{0, 1, ..., q-1\}$. Let Y be the received symbol with the output alphabet $\mathcal{Y} = \left\{ \mathcal{X} \bigcup_{x=0}^{q-1} \left\{ ?_x^{(i)} \right\}_{i=1}^{i_{\max}} \right\}$, where each super-symbol $?_x^{(i)}$ (for $x = 0, 1, ..., q-1$) consists of a set of size M that contains the symbol x and $M-1$ other symbols, taken from $\mathcal{X} \setminus \{x\}$. Let's denote by $\ell(n,k)$ the binomial coefficient $\begin{pmatrix} n \\ k \end{pmatrix}$. Clearly, there are $i_{\max} = \ell(q-1, M-1)$ possible super-symbols for each x.

The transition probabilities governing the QPEC are as follows:

$$\Pr(Y = y | X = x) = \begin{cases} 1-\varepsilon, & y = x \\ \varepsilon / i_{\max}, & y = ?_x^{(i)} \end{cases} \quad (1)$$

for $i = 1, 2, ..., i_{\max}$, where $0 \le \varepsilon \le 1$ is the (partial) erasure probability. That is, the output of the channel can be either a *symbol*, with probability $1-\varepsilon$ (corresponding to a non-erasure event), or a *set of M symbols*, with probability ε (corresponding to a partial erasure event). As an example, assume that $q = 4, M = 2$, and the symbol 0 was transmitted. Then we have $?_x^{(1)} = \{0,1\}, ?_x^{(2)} = \{0,2\}$ and $?_x^{(3)} = \{0,3\}$, where each is received with probability $\varepsilon/3$ and 0 is received with probability $1-\varepsilon$.

Note that for $M = q$ we get the q-ary erasure channel (QEC), the common generalization of the BEC to $q > 2$. In our analysis, we will use the arithmetic of the finite field GF(q), such that q will be a prime or a prime power, and the symbols in $\mathcal{X}$ will be assumed to be the elements of GF(q).

B. Capacity

Denote $p_k = \Pr(X = k)$, for $k = 0, 1, ..., q-1$, to be the input distribution to the channel. The channel capacity C is:

$$C = \max_{\{p_k\}_{k=0}^{q-1}} I(X;Y) = \max_{\{p_k\}_{k=0}^{q-1}} (H(Y) - H(Y|X)), \quad (2)$$

where $I(X;Y)$ is the mutual information between the input X and the output Y, and $H(Y)$, $H(Y|X)$ are the entropy of Y and the conditional entropy of Y given X, respectively. Similarly to the case of the BEC, the mutual information is maximized under the uniform distribution of the input.

Theorem 1. (Capacity achieving input distribution for the QPEC) *Assume a QPEC channel with an input probability distribution* $\{p_k\}_{k=0}^{q-1}$*. Then the capacity is achieved for the uniform distribution of the input, and we have:*

$$C(QPEC) = 1 - \varepsilon \log_q M \tag{3}$$

measured in q-ary symbols per channel use.

Proof. The proof is based on the use of Lagrange multipliers and is omitted due to space limit. A detailed proof appears in [6]. □

Note that the capacity C for the QPEC is in agreement with the capacity of the QEC ($M = q$) and in particular with the capacity of the BEC ($M = q = 2$).

III. Message Passing Algorithm for the QPEC

A GF(q) LDPC $[n, k]$ code is defined in a similar way to its binary counterpart, by a sparse parity-check matrix, or equivalently by a Tanner graph [7]. This graph is bipartite, with n variable (left) nodes, which correspond to symbols of the codeword, and $n-k$ check (right) nodes, which correspond to parity check equations. The codeword symbols and the labels on the edges of the graph are taken from GF(q). For ease of presentation we will concentrate here on *regular* LDPC codes, having a constant check node degree d_c and a constant variable node degree d_v.

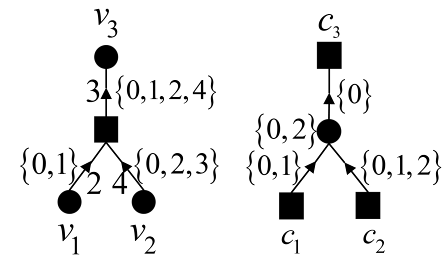
In the graph, each check node c_j is connected, by edges, to variable nodes $v_i, i \in N(j)$, where $N(j)$ denotes the set of nodes adjacent to node i. The parity check equation induced by c_j is satisfied when $\sum_{i \in N(j)} h_{ij} v_i = 0$, where h_{ij} is the label on the edge connecting variable node i to check node j.

The following decoder for q-ary LDPC codes over the QPEC is a variation of the standard message passing/belief propagation algorithm over a Tanner graph to match the partial information exchanged in decoding. For this decoder, the beliefs exchanged in the decoding process are *sets of symbols*, rather than probabilities. We have two types of messages: *check to variable* (CTV) messages, and *variable to check* (VTC) messages, denoted by $c_{j \to i}$ and $v_{i \to j}$, respectively.

At iteration $l = 0$, channel information is sent from variable to check nodes: erased nodes send sets of size M, and non-erased ones send sets of size 1 (containing the correct symbol). In the next iterations, we have the following messages:

1) **Check to variable (CTV)**.
First, we define for each $i' \in N(j) \setminus i$ the following set:

$$X_{i'}^{(l)} = \left\{ -\frac{h_{i'j} \cdot x_{i'}}{h_{ij}} : x_{i'} \in v_{i' \to j}^{(l-1)} \right\}, \quad l \geq 1. \tag{4}$$



(a) CTV message (b) VTC message

Fig. 1: Message passing: examples (GF(5))

Then we have:

$$\begin{aligned} c_{j \to i}^{(l)} &= \sum_{i' \in N(j) \setminus i} X_{i'}^{(l)} \\ &\triangleq \left\{ \sum_{i' \in N(j) \setminus i} a_{i'} : a_{i'} \in X_{i'}^{(l)} \right\}, \quad l \geq 1 \end{aligned} \tag{5}$$

where the calculations are carried over GF(q). In words, the message (5) consists of all possible assignments of the variable node i, such that the parity equation involving the variable nodes in the set $N(j)$ is satisfied. An example for a CTV message is given in Figure 1a. In this example (over GF(5)), the variable nodes v_1, v_2 and v_3 are connected to the same check node, with edges having the labels $2, 4$ and 3. v_1 is known to be either 0 or 1, and v_2 is known to be either 0, 2 or 3. Therefore, the outgoing message is consisted of all possible outcomes of the expression $(2v_1 + 4v_2)/(-3)$.

2) **Variable to check (VTC)**.

$$v_{i \to j}^{(l)} = v_{i \to j}^{(0)} \bigcap \left(\bigcap_{j' \in N(i) \setminus j} c_{j' \to i}^{(l)} \right), \quad l \geq 1 \tag{6}$$

where $v_{i \to j}^{(0)}$ is the output from the channel information for variable node i, which is passed at iteration 0. The resulting message is simply the intersection of the incoming messages and the channel information. An example for a VTC message is given in Figure 1b.

In practice, the decoder stops after a finite number of iterations. The decoding is declared successful if the size of all VTC messages is 1.

IV. Decoding Analysis through Density Evolution

The density evolution method proposed in [8] is an analytical tool for evaluating the asymptotic performance of LDPC codes under message-passing decoding. Note that in our case the *all-zero codeword assumption* [8] holds, due to the symmetry of the QPEC and the message passing algorithm.

The key idea we use for analyzing the densities is to track the probability distribution on the *sizes* of the messages, leading to just q entries in the distribution, instead of $2^q - 1$. This approach is a more natural one in our case, since a decoding failure may occur when a VTC message has size larger than 1, independent of the exact content of the message.

A. Density-evolution equations

In this part, we present the density evolution equations corresponding to BP decoding for QPEC, assuming that the LDPC graph was drawn at random. In the following, $\mathbf{w}^{(l)}$ is a probability vector, where $w_m^{(l)}$ $(m = 1, 2, ..., q)$ denotes the probability that a CTV message at iteration l is of size m. The probability vector $\mathbf{z}^{(l)}$ is defined for VTC messages in a similar manner.

The following density-evolution equations are based on the following idea. For each possible set of sizes of incoming messages, its probability is calculated by multiplying the probability of each incoming message size. This probability is then multiplied by the probability that the outgoing message will be of size m, given the sizes of the incoming messages. We get:

1) **CTV messages:**

$$w_m^{(l)} = \sum_{\substack{\{|S_j|\}_{j=1}^{d_c-1} : \\ |S_j| \le M}} \left(\prod_{j=1}^{d_c-1} z_{|S_j|}^{(l-1)} \right) \cdot P_m \left(\{|S_j|\}_{j=1}^{d_c-1} \right). \tag{7}$$

The initial conditions are $z_1^{(0)} = 1 - \varepsilon, z_M^{(0)} = \varepsilon$ and $z_m^{(0)} = 0$ for $m \neq 1, M$. P_m denotes the probability that a CTV message is of size m, given the sizes of the incoming VTC messages, $\{|S_j|\}_{j=1}^{d_c-1}$.

2) **VTC messages:**

$$z_m^{(l)} = \delta[m-1] \cdot (1-\varepsilon) + \varepsilon \sum_{\substack{\{|S_j|\}_{j=1}^{d_v-1} : \\ |S_j| \le q}} \left(\prod_{j=1}^{d_v-1} w_{|S_j|}^{(l)} \right) \cdot Q_m \left(\left\{ \{|S_j|\}_{j=1}^{d_v-1}, M \right\} \right). \tag{8}$$

$\delta[m]$ denotes the discrete Dirac delta function. Q_m denotes the probability that a VTC message is of size m, given the sizes of the incoming CTV messages, $\{|S_j|\}_{j=1}^{d_v-1}$, and the size M of the partially-erased variable node.

Finding the exact P_m as a function of the incoming message sizes is a hard problem, as we will see in Section IV-B, where several bounds over P_m will be given. We also give two approximation models for P_m in Section V. On the other hand, an exact formula for Q_m will be provided in Section IV-C. Note that regardless of the exact behaviour of P_m, the density-evolution equation of the BEC [3] (hence QEC) can be derived from Equations (7) and (8) by setting $M = q$ [6].

B. Equivalent formulation for P_m and bounds

Assume that we have $K = d_c - 1$ subsets of GF(q), $\{S_j\}_{j=1}^K$. Their *sumset*, denoted $\sum_{j=1}^K S_j$, is defined as follows:

$$\sum_{j=1}^K S_j \triangleq \left\{ \sum_{j=1}^K s_j : s_j \in S_j \right\}. \tag{9}$$

That is, the sumset of the subsets $\{S_j\}_{j=1}^K$ is defined to be the set of all sums (using GF(q) arithmetic) of elements taken from the subsets. When the labels h_{ij} of the graph are chosen at random, the CTV message of Equation (5) can be considered as a sumset of random subsets of GF(q). Noting that, P_m is equivalent to the probability that a sumset of random subsets is of size m, when the sizes of the subsets are known.

Finding the number of elements within the sumset as a function of $\{|S_j|\}_{j=1}^K$ is an open problem in additive combinatorics (see e.g. [9]). This stems from the structure of the field, where a symbol in a field can be obtained by multiple combinations of sums of symbols. In the following, we provide bounds on the size of the sumset.

Lemma 2. *Consider K non-empty subsets of GF(q), $\{S_j\}_{j=1}^K$. Then:*

$$\max_j |S_j| \le \left| \sum_{j=1}^K S_j \right| \le \min \left(q, \prod_{j=1}^K |S_j| \right). \tag{10}$$

Proof. The proof is immediate using basic properties of fields. See [6] for details. □

We can improve the bounds (10), by using the following two theorems and their corollary.

Theorem 3. (Cauchy-Davenport Theorem [10]) *Consider the finite field GF(p), p prime, where A and B are two non-empty subsets of GF(p). Then:*

$$|A + B| = \{ a + b | \, a \in A, b \in B \} \ge \min (p, |A| + |B| - 1).$$

Theorem 4. (Károlyi's Theorem for Finite Groups [11]) *Let A and B are two non-empty subsets of a finite group G. Denote by $p(G)$ the smallest prime factor of $|G|$. Then:*

$$|A + B| \ge \min (p(G), |A| + |B| - 1).$$

Corollary 5. *Assume a finite field GF(q), where $q = p^s$, p is prime and s is a positive integer. Then:*

$$\max \left(\max_j |S_j|, \min \left(p, \sum_{j=1}^K |S_j| - K + 1 \right) \right) \le \left| \sum_{j=1}^K S_j \right| \le \min \left(q, \prod_{j=1}^K |S_j| \right). \tag{11}$$

Proof. This corollary is proved by Lemma 2 and Theorems 3 and 4, followed by induction on the number of subsets. □

We will denote by B_L and B_U the lower and upper bounds of (11), respectively. We have the following sufficient condition for attaining the maximum size, q, of the sumset.

Proposition 6. (Sufficient condition for $\left| \sum_{j=1}^K S_j \right| = q$) *Consider K non-empty subsets of GF(q), $\{S_j\}_{j=1}^K$. If there is a pair of sets $S_a, S_b \in \{S_j\}_{j=1}^K$ $(a \neq b)$ such that $|S_a| + |S_b| > q$, then $\left| \sum_{j=1}^K S_j \right| = q$.*

Proof. This is a known result from the theory of finite groups. See [6] for a detailed proof. □

For later use, we say that the *q-condition* holds if the condition of Proposition 6 is satisfied. Using the bounds B_L and B_U and the q-condition, we get the following bounds (in terms of the size of the sumset) for P_m:

$$P_m^{(\max)} = \begin{cases} \delta[m-q], & \text{if the } q\text{-condition holds} \\ \delta[m-B_U], & \text{otherwise} \end{cases} \quad (12)$$

$$P_m^{(\min)} = \begin{cases} \delta[m-q], & \text{if the } q\text{-condition holds} \\ \delta[m-B_L], & \text{otherwise} \end{cases} \quad (13)$$

Using the above $P_m^{(\max)}$ resp. $P_m^{(\min)}$ in (7) will give a lower resp. upper bound on the decoding threshold.

C. Equivalent formulation and formula for Q_m

When the labels h_{ij} are chosen at random, Q_m is equivalent to the probability that the intersection of $J = d_v$ random GF(q) subsets with sizes $\left\{ \{|S_j|\}_{j=1}^{d_v-1}, M \right\}$ (M corresponds to the size of the set provided by the channel information) is exactly m. We begin with the following lemma.

Lemma 7. *Assume that $\{S_j\}_{j=1}^{J}$ are subsets of a set with q elements, with given sizes $\{|S_j|\}_{j=1}^{J}$, where $\mu \triangleq \min_j |S_j|$. Then, the number of ways to get an intersection of size m ($m = 0, 1, ..., \mu$) between the subsets is:*

$$I_m\left(\{|S_j|\}_{j=1}^{J}; q\right) = \sum_{i=0}^{\mu-m} (-1)^i \cdot v\left(\{|S_j|\}_{j=1}^{J}, m+i\right) \cdot \ell(m+i, m), \quad (14)$$

where

$$v\left(\{|S_j|\}_{j=1}^{J}, l\right) = \ell(q, l) \cdot \prod_{j=1}^{J} \ell(q-l, |S_j| - l). \quad (15)$$

Proof. This can be proved using the inclusion-exclusion principle. See [6] for details. □

We are now ready to provide an exact formula for Q_m.

Theorem 8. *Assume that we are given the sizes $\{|S_j|\}_{j=1}^{J}$ of J subsets of GF(q), where each subset contains the symbol 0 (as can be assumed due to the all-zero codeword assumption). Then, the probability for an intersection of size m ($m = 1, 2, ..., \mu = \min(\{|S_j|\}_{j=1}^{J})$ between the sets is:*

$$Q_m\left(\{|S_j|\}_{j=1}^{J}; q\right) = \begin{cases} \dfrac{I_{m-1}\left(\{|S_j|-1\}_{j=1}^{J}; q-1\right)}{\prod_{j=1}^{J} \ell(q-1, |S_j|-1)}, & \text{if } \mu > 1 \\ \delta[m-1], & \text{otherwise} \end{cases} \quad (16)$$

where I_m is defined in Equation (14).

Proof. This is a result of Lemma 7. We use I_{m-1} instead of I_m since the symbol 0 belongs to each set. Moreover, instead of q possible elements to choose from, we have only $q-1$ possible ones, since 0 is already taken. Finally, we normalize by the number of subsets with sizes $|S_j| - 1$ taken from a set of $q-1$ elements. □

V. Approximation models for P_m

So far, we provided an exact formula for Q_m and bounds for P_m. An exact expression for P_m is likely difficult to find. In this section, we discuss appropriate models for approximating $P_m\left(\{|S_j|\}_{j=1}^{K}\right)$ from Equation (7) (where $K = d_c - 1$). We begin with a simple *balls-and-bins model*, and later refine it with a tighter model we term as the *union model*.

A. The balls and bins model

In the balls and bins model [12], there is a set of balls and a set of bins. Each bin is assumed to be picked independently and uniformly at random for each ball. We consider each sum within the sumset (9) as a ball, and the q elements of GF(q) as bins. Our aim is to calculate the probability that $N = \prod_{j=1}^{K} |S_j|$ balls (sums) are assigned to m bins ($m = 1, 2, ..., q$).

For this task, we provide a formulation of the balls and bins model as a *Markov process*. First, define:

$$T_m\left(\{|S_j|\}_{j=1}^{K}\right) = \frac{I_m\left(\{|S_j|\}_{j=1}^{K}\right)}{\prod_{j=1}^{K} \ell(q, |S_j|)}, \quad (17)$$

using I_m from Equation (14) (q was omitted for convenience). T_m is the probability that the intersection of randomly chosen subsets $\{S_j\}_{j=1}^{K}$ of a set of size q is of size m.

We define the $q+1$ possible *states* ($i = 0, 1, ..., q$) in the balls and bins model as the number of non-empty bins. We now write the Markov matrix associated with this model. The (i, j) entry in this matrix is the transition probability from a state with i non-empty bins to a state with j non-empty bins:

$$(\mathbf{\Gamma}_{\text{balls}})_{i,j} = T_{1+i-j}(\{1, i\}), \quad (18)$$

with indices i, j ranging from 0 to q. Simple calculations show that $(\mathbf{\Gamma}_{\text{balls}})_{i,i} = \frac{i}{q}$ and $(\mathbf{\Gamma}_{\text{balls}})_{i,i+1} = 1 - \frac{i}{q}$ while the remaining elements of $\mathbf{\Gamma}_{\text{balls}}$ are zeros.

Define the probability vector $\mathbf{g}^{(l)} = \left(g_0^{(l)}, g_1^{(l)}, ..., g_q^{(l)}\right)$ over the states defined by $\mathbf{\Gamma}_{\text{balls}}$. Clearly, $\mathbf{g}^{(0)} = (1, 0, ..., 0)$. Now, using the Markov property, we get:

$$\mathbf{g}^{(l)} = \mathbf{g}^{(0)} \mathbf{\Gamma}_{\text{balls}}^{l}. \quad (19)$$

Taking into account the lower bound B_L and the *q-condition*, we get the following approximation for P_m:

$$P_m^{(\text{balls})} = \begin{cases} 0, & \text{if } m < B_L \\ \delta[m-q], & \text{if the } q\text{-condition holds} \\ \dfrac{g_m^{(N)}}{\sum_{i=B_L}^{q} g_i^{(N)}}, & \text{otherwise} \end{cases} \quad (20)$$

B. The union model

According to the balls and bins model, each ball (sum) is assigned to a bin independently. However, it is clear (according to Lemma 2) that the sums that appear within the sumset (9) can be divided into N/κ sets of κ *distinct* elements each, where $\kappa \triangleq \max_j |S_j|$. To take this into account, we next introduce the *union model*.

According to the union model, the probability for a sumset of size m is modelled as the probability that m bins are non-empty after N balls are assigned uniformly at random to the q bins *set-by-set*, rather than one-by-one, where each set of size κ is assigned to κ *distinct* bins. Similarly to the balls and bins model, the union model can also be formulated in terms of a Markov process, using the matrix:

$$\left(\mathbf{\Gamma}_{\text{union}}\right)_{i,j} = T_{\kappa+i-j}\left(\{\kappa, i\}\right). \tag{21}$$

Define the probability vector $\mathbf{u}^{(l)} = \left(u_0^{(l)}, u_1^{(l)}, ..., u_q^{(l)}\right)$ over the states defined by $\mathbf{\Gamma}_{\text{union}}$. Clearly, $\mathbf{u}^{(0)} = (1, 0, ..., 0)$. We have the following relation:

$$\mathbf{u}^{(l)} = \mathbf{u}^{(0)}\mathbf{\Gamma}_{\text{union}}^{l}. \tag{22}$$

Finally, we get the following approximation for P_m:

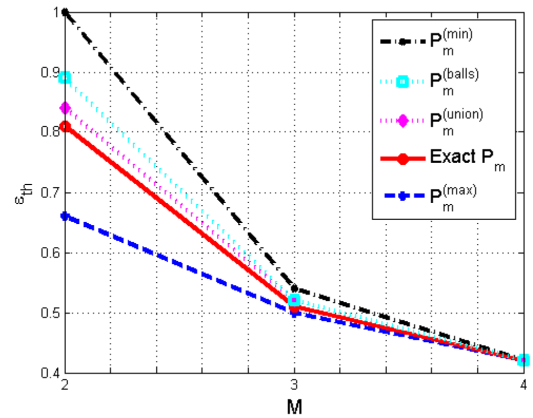
$$P_m^{(\text{union})} = \begin{cases} 0, & \text{if } m < B_L \\ \delta\left[m-q\right], & \text{if the } q\text{-condition holds} \\ \dfrac{u_m^{(N/\kappa)}}{\sum\limits_{i=B_L}^{q} u_i^{(N/\kappa)}}, & \text{otherwise} \end{cases} \tag{23}$$

C. Decoding threshold estimation

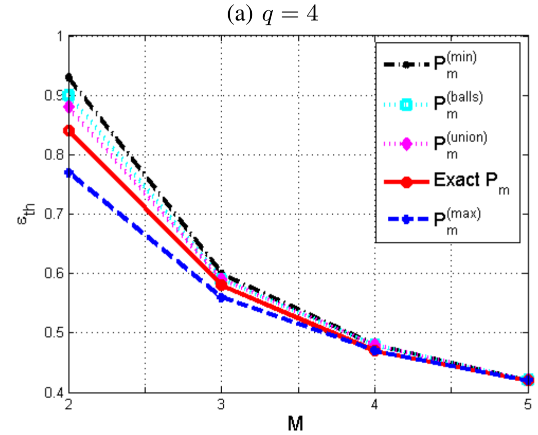
As in the case of BEC/QEC, we have a threshold phenomenon [3] for the QPEC. The threshold is the maximal allowed ε such that the probability of decoding error vanishes after sufficient number of iterations. It can be thought of as the maximal allowed fraction of partially known (up to M levels) symbols in a q-level flash memory.

In Figure 2, we provide the threshold (denoted ε_{th}) for a regular $(3, 6)$ LDPC code, calculated using the density evolution equations (7) and (8) for $q = 4$ and $q = 5$. The exact P_m was calculated numerically (by running over all possible assignments of sets).

Both the balls and bins model and the union model appear to serve as an upper bound for ε_{th}, with the union model being tighter. Moreover, both models seem to provide better estimate of the threshold as M approaches q. As mentioned in Section IV, the threshold for $M = q$ is the threshold of the BEC.



(a) $q = 4$



(b) $q = 5$

Fig. 2: The threshold ε_{th} for a regular $(3, 6)$ LDPC code

VI. Conclusion

In this paper, we defined a new channel - the QPEC - motivated by multilevel NVMs. We provided an appropriate belief propagation decoder for this channel when used with LDPC codes, with the corresponding density evolution equations. We developed approximation models for these equations, since their exact analysis is closely related to an open problem in additive combinatorics. The results show the importance of these models, which appear to provide good approximations.

VII. Acknowledgement

This work was supported in part by the Israel Ministry of Science and Technology, by an EU CIG grant, and by a Viterbi faculty fellowship.

References

[1] Y. Cassuto, M. Schwartz, V. Bohossian, and J. Bruck, "Codes for asymmetric limited-magnitude errors with application to multi-level flash memories," *IEEE Transactions on Information theory*, vol. 56, no. 4, pp. 1582–1595, 2010.

[2] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[3] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[4] J. Wang, T. Courtade, H. Shankar, and R. Wesel, "Soft information for LDPC decoding in flash: Mutual-information optimized quantization," in *IEEE GLOBECOM 2011*, pp. 1–6.

[5] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete-memoryless channels," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 549–583, 2006.

[6] R. Cohen and Y. Cassuto, "LDPC codes for partial-erasure channels in multi-level memories," *arXiv:1401.4799 [cs.IT]*, 2014.

[7] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory* , vol. IT-27, pp. 533–547, 1981.

[8] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 599–618, 2001.

[9] E. Croot and V. F. Lev, "Open problems in additive combinatorics," Department of Mathematics, Georgia Institute of Technology, Tech. Rep., 2011.

[10] T. C. Tao and V. H. Vu, *Additive Combinatorics*. Cambridge University Press, 2006.

[11] G. Károlyi, "The Cauchy-Davenport theorem in group extensions," *L'Enseignement Mathématique 51*, 2005.

[12] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.