# Growth Rate of the Weight Distribution of Doubly-Generalized LDPC Codes: General Case and Efficient Evaluation

Mark F. Flanagan, Enrico Paolini, Marco Chiani, and Marc P. C. Fossorier

*Abstract*—The growth rate of the weight distribution of irregular doubly-generalized LDPC (D-GLDPC) codes is developed and in the process, a new efficient numerical technique for its evaluation is presented. The solution involves simultaneous solution of a $4 \times 4$ system of polynomial equations. This represents the first efficient numerical technique for *exact* evaluation of the growth rate, even for LDPC codes. The technique is applied to two example D-GLDPC code ensembles.

## I. INTRODUCTION

Recently, the design and analysis of coding schemes representing generalizations of Gallager's low-density parity-check (LDPC) codes [1] has gained increasing attention. This interest is motivated above all by the potential capability of these coding schemes to offer a better compromise between waterfall and error floor performance than is currently offered by state-of-the-art LDPC codes.

In the Tanner graph of an LDPC code, any degree-$q$ variable node (VN) may be interpreted as a length-$q$ repetition code, i.e., as a $(q, 1)$ linear block code. Similarly, any degree-$s$ check node (CN) may be interpreted as a length-$s$ single parity-check (SPC) code, i.e., as a $(s, s - 1)$ linear block code. The first proposal of a class of linear block codes generalizing LDPC codes may be found in [2], where it was suggested to replace each CN of a regular LDPC code with a generic linear block code, to enhance the overall minimum distance. The corresponding coding scheme is known as a regular generalized LDPC (GLDPC) code, or Tanner code, and a CN that is not a SPC code as a generalized CN. More recently, irregular GLDPC codes were considered (see for instance [3]). For such codes, the VNs exhibit different degrees and the CN set is composed of a mixture of different linear block codes.

A further generalization step is represented by doubly-generalized LDPC (D-GLDPC) codes [4]. In a D-GLDPC code, not only the CNs but also the VNs may be represented by generic linear block codes. The VNs which are not repetition codes are called generalized VNs. The main motivation for introducing generalized VNs is to overcome some problems connected with the use of generalized CNs, such as an overall

code rate loss which makes GLDPC codes interesting mainly for low code rate applications, and a loss in terms of decoding threshold (for a discussion on drawbacks of generalized CNs and on beneficial effects of generalized VNs we refer to [5] and [6], respectively).

A useful tool for analysis and design of LDPC codes and their generalizations is represented by the growth rate of the weight distribution, or equivalently, the asymptotic weight enumerating function (WEF). The growth rate of the weight distribution was introduced in [1] to show that the minimum distance of a randomly generated regular LDPC code with a VN degree of at least three is a linear function of the codeword length with high probability. The same approach was taken in [7] and [8] to obtain related results on the minimum distance of subclasses of Tanner codes.

The growth rate of the weight distribution has been subsequently investigated for unstructured ensembles of irregular LDPC codes. Works in this area are [9], [10], [11], [12]. In particular, in [12] a technique for approximate evaluation of the growth rate of any (eventually expurgated) irregular LDPC ensemble has been developed, based on Hayman's formula. Asymptotic weight enumerators of ensembles of irregular LDPC codes based on protographs and on multiple edge types have been derived in [13] and [14], respectively. The approach proposed in [13] has then been extended to protograph GLDPC codes and to protograph D-GLDPC codes in [15] and [16], respectively. In [17], the authors presented a compact formula for the growth rate of general unstructured irregular D-GLDPC code ensembles for the specific case of *small* weight codewords.

In this paper, an analytical expression for the growth rate of the weight distribution of a general unstructured irregular ensemble of D-GLDPC codes is developed. As opposed to the formula developed in [17], the proposed expression holds for any codeword weight. The present work also extends to the fully-irregular case an expression for the growth rate obtained in [18] assuming a CN set composed of linear block codes all of the same type. In the process of this development, we obtain an efficient evaluation tool for computing the growth rate exactly. This tool always requires the solution of a $(4 \times 4)$ polynomial system of equations, *regardless* of the number of VN types and CN types in the D-GLDPC ensemble. As shown through numerical examples, the proposed tool allows to obtain a precise plot of the growth rate with a low computational effort. For the case of irregular LDPC codes, a technique for numerical evaluation of the growth rate of

M. F. Flanagan is with the School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Belfield, Dublin 4, Ireland (e-mail:mark.flanagan@ieee.org).

E. Paolini and M. Chiani are with DEIS, University of Bologna, Via Venezia 52, 47023 Cesena (FC), Cesena, Italy (e-mail:e.paolini@unibo.it, marco.chiani@unibo.it).

M. P. C. Fossorier is with ETIS ENSEA, UCP, CNRS UMR-8051, 6 avenue du Ponceau, 95014 Cergy Pontoise, France (e-mail: mfossorier@ieee.org).

the weight distribution was given in [12]; in contrast to the technique developed in this paper, the method of [12] provided an *approximate* numerical solution for the growth rate; it is also more computationally complex than that proposed in the present work.

## II. PRELIMINARIES AND NOTATION

We define a D-GLDPC code ensemble $\mathcal{M}_n$ as follows, where $n$ denotes the number of VNs. There are $n_c$ different CN types $t \in I_c = \{1, 2, \cdots, n_c\}$, and $n_v$ different VN types $t \in I_v = \{1, 2, \cdots, n_v\}$. For each CN type $t \in I_c$, we denote by $h_t$, $s_t$ and $r_t$ the CN dimension, length and minimum distance, respectively. For each VN type $t \in I_v$, we denote by $k_t$, $q_t$ and $p_t$ the VN dimension, length and minimum distance, respectively. For $t \in I_c$, $\rho_t$ denotes the fraction of edges connected to CNs of type $t$. Similarly, for $t \in I_v$, $\lambda_t$ denotes the fraction of edges connected to VNs of type $t$. Note that all of these variables are independent of $n$.

The polynomials $\rho(x)$ and $\lambda(x)$ are defined by $\rho(x) \triangleq \sum_{t \in I_c} \rho_t x^{s_t - 1}$ and $\lambda(x) \triangleq \sum_{t \in I_v} \lambda_t x^{q_t - 1}$. If $E$ denotes the number of edges in the Tanner graph, the number of CNs of type $t \in I_c$ is then given by $E\rho_t/s_t$, and the number of VNs of type $t \in I_v$ is then given by $E\lambda_t/q_t$. Denoting as usual $\int_0^1 \rho(x)\,\mathrm{d}x$ and $\int_0^1 \lambda(x)\,\mathrm{d}x$ by $\int \rho$ and $\int \lambda$ respectively, we see that the number of edges in the Tanner graph is given by $E = n/\int \lambda$ and the number of CNs is given by $m = E \int \rho$. Therefore, the fraction of CNs of type $t \in I_c$ and the fraction of VNs of type $t \in I_v$ are given by

$$\gamma_t = \frac{\rho_t}{s_t \int \rho} \quad \text{and} \quad \delta_t = \frac{\lambda_t}{q_t \int \lambda} \tag{1}$$

respectively. Also the length of any D-GLDPC codeword in the ensemble is given by

$$N = \sum_{t \in I_v} \left( \frac{E\lambda_t}{q_t} \right) k_t = \frac{n}{\int \lambda} \sum_{t \in I_v} \frac{\lambda_t k_t}{q_t} . \tag{2}$$

Note that this is a linear function of $n$. Similarly, the total number of parity-check equations for any D-GLDPC code in the ensemble is given by $M = \frac{m}{\int \rho} \sum_{t \in I_c} \frac{\rho_t(s_t - h_t)}{s_t}$. A member of the ensemble $\mathcal{M}_n$ then corresponds to a permutation on the $E$ edges connecting CNs to VNs.

The WEF for CN type $t \in I_c$ is given by $A^{(t)}(z) = 1 + \sum_{u=r_t}^{s_t} A_u^{(t)} z^u$. Here $A_u^{(t)} \geq 0$ denotes the number of weight-$u$ codewords for CNs of type $t$. The input-output weight enumerating function (IO-WEF) for VN type $t \in I_v$ is given by $B^{(t)}(x, y) = 1 + \sum_{u=1}^{k_t} \sum_{v=p_t}^{q_t} B_{u,v}^{(t)} x^u y^v$. Here $B_{u,v}^{(t)} \geq 0$ denotes the number of weight-$v$ codewords generated by input words of weight $u$, for VNs of type $t$. Also, $B_2^{(t)}$ is the total number of weight-2 codewords for VNs of type $t$.

If there exist CNs and VNs with minimum distance equal to 2, and define the (positive) parameters

$$C = 2 \sum_{t \,:\, r_t = 2} \frac{\rho_t A_2^{(t)}}{s_t} ; \quad V = 2 \sum_{t \,:\, p_t = 2} \frac{\lambda_t B_2^{(t)}}{q_t} . \tag{3}$$

The design rate of any D-GLDPC ensemble is given by

$$R = 1 - \frac{\sum_{t \in I_c} \rho_t(1 - R_t)}{\sum_{t \in I_v} \lambda_t R_t} \tag{4}$$

where for $t \in I_c$ (resp. $t \in I_v$) $R_t$ is the local code rate of a type-$t$ CN (resp. VN).

Throughout this paper, the notation $e = \exp(1)$ denotes Napier's number, all the logarithms are assumed to have base $e$ and for $0 < x < 1$ the notation $h(x) = -x\log(x) - (1 - x)\log(1 - x)$ denotes the binary entropy function.

## III. GROWTH RATE OF THE WEIGHT DISTRIBUTION OF GENERAL IRREGULAR D-GLDPC CODE ENSEMBLES

The growth rate of the weight distribution of the irregular D-GLDPC ensemble sequence $\{\mathcal{M}_n\}$ is defined by

$$G(\alpha) \triangleq \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}_{\mathcal{M}_n} [N_{\alpha n}] \tag{5}$$

where $\mathbb{E}_{\mathcal{M}_n}$ denotes the expectation operator over the ensemble $\mathcal{M}_n$, and $N_w$ denotes the number of codewords of weight $w$ of a randomly chosen D-GLDPC code in the ensemble $\mathcal{M}_n$. The limit in (5) assumes the inclusion of only those positive integers $n$ for which $\alpha n \in \mathbb{Z}$ and $\mathbb{E}_{\mathcal{M}_n}[N_{\alpha n}]$ is positive. Note that the argument of the growth rate function $G(\alpha)$ is equal to the ratio of D-GLDPC codeword length to the number of VNs; by (2), this captures the behaviour of codewords linear in the block length, as in [12] for the LDPC case.

A D-GLDPC ensemble is said to be *asymptotically good* if and only if $\alpha^* \triangleq \inf\{\alpha > 0 \mid G(\alpha) \geq 0\} > 0$. The parameter $\alpha^*$ is called the *ensemble relative minimum distance*. In [19], it was shown that a D-GLDPC ensemble is always asymptotically good if there exist no CNs or VNs with minimum distance 2 while, if the exist both CNs and VNs with minimum distance 2, the ensemble is asymptotically good if and only if $C \cdot V < 1$, where $C$ and $V$ are given by (3).

Note that using (2), we may also define the growth rate with respect to the number of D-GLDPC code bits $N$ as follows:

$$H(\gamma) \triangleq \lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}_{\mathcal{M}_n} [N_{\gamma N}] . \tag{6}$$

It is straightforward to show that

$$H(\gamma) = \frac{G(\gamma y)}{y} \tag{7}$$

where

$$y = \frac{1}{\int \lambda} \sum_{t \in I_v} \frac{\lambda_t k_t}{q_t} .$$

In this section, we formulate an expression of the growth rate for an irregular D-GLDPC ensemble $\mathcal{M}_n$ over a wider range of $\alpha$ than was considered in [17], [19] (where the case $\alpha \to 0$ was analyzed).

The following theorem constitutes our main result.

*Theorem 1:* The growth rate of the weight distribution of the irregular D-GLDPC ensemble sequence $\{\mathcal{M}_n\}$ is given by

$$G(\alpha) = \sum_{t \in I_v} \delta_t \log B^{(t)}(x_0, y_0) - \alpha \log x_0$$

$$+ \left( \frac{\int \rho}{\int \lambda} \right) \sum_{s \in I_c} \gamma_s \log A^{(s)}(z_0) + \frac{\log \left( 1 - \beta \int \lambda \right)}{\int \lambda} \tag{8}$$

where $x_0$, $y_0$, $z_0$ and $\beta$ are the unique positive real solutions to the $4 \times 4$ system of polynomial equations[1]

$$z_0 \left( \frac{\int \rho}{\int \lambda} \right) \sum_{t \in I_c} \gamma_t \frac{\frac{\mathrm{d}A^{(t)}}{\mathrm{d}z}(z_0)}{A^{(t)}(z_0)} = \beta , \tag{9}$$

$$x_0 \sum_{t \in I_v} \delta_t \frac{\frac{\partial B^{(t)}}{\partial x}(x_0, y_0)}{B^{(t)}(x_0, y_0)} = \alpha , \tag{10}$$

$$y_0 \sum_{t \in I_v} \delta_t \frac{\frac{\partial B^{(t)}}{\partial y}(x_0, y_0)}{B^{(t)}(x_0, y_0)} = \beta , \tag{11}$$

and

$$\left( \beta \int \lambda \right) (1 + y_0 z_0) = y_0 z_0 . \tag{12}$$

The theorem is proved in Section IV.

## IV. Proof of the Main Result

In this section we prove Theorem 1. The proof uses the concepts of *assignment* and *split assignment*, defined next.

*Definition 1:* An *assignment* is a subset of the edges of the Tanner graph. An assignment is said to have *weight* $k$ if it has $k$ elements. An assignment is said to be *check-valid* if the following condition holds: supposing that each edge of the assignment carries a 1 and each of the other edges carries a 0, each CN recognizes a valid local codeword.

*Definition 2:* A *split assignment* is an assignment, together with a subset of the D-GLDPC code bits (called a *codeword assignment*). A split assignment is said to have *split weight* $(u, v)$ if its assignment has weight $v$ and its codeword assignment has $u$ elements. A split assignment is said to be *check-valid* if its assignment is check-valid. A split assignment is said to be *variable-valid* if the following condition holds: supposing that each edge of its assignment carries a 1 and each of the other edges carries a 0, and supposing that each D-GLDPC code bit in the codeword assigment is set to 1 and each of the other code bits is set to 0, each VN recognizes a local input word and the corresponding valid local codeword.

For ease of presentation, the proof is broken into two parts.

### A. Number of check-valid assignments of weight $\delta m$

First we derive an expression, valid asymptotically, for the number of check-valid assignments of weight $\delta m$. For each $t \in I_c$, let $\epsilon_t m$ denote the portion of the total weight $\delta m$ apportioned to CNs of type $t$. Then $\epsilon_t \geq 0$ for each $t \in I_c$, and $\sum_{t \in I_c} \epsilon_t = \delta$. Also denote $\boldsymbol{\epsilon} = (\epsilon_1 \ \epsilon_2 \ \cdots \ \epsilon_{n_c})$.

Consider the set of $\gamma_t m$ CNs of a particular type $t \in I_c$, where $\gamma_t$ is given by (1). Using generating functions, the number of check-valid assignments (over these CNs) of weight $\epsilon_t m$ is given by

$$N_{c,t}^{(\gamma_t m)}(\epsilon_t m) = \mathrm{Coeff} \left[ \left( A^{(t)}(x) \right)^{\gamma_t m}, x^{\epsilon_t m} \right]$$

where $\mathrm{Coeff}\left[ p(x), x^c \right]$ denotes the coefficient of $x^c$ in the polynomial $p(x)$. We next make use of the following result, which is a special case of [12, Corollary 16]:

*Lemma 1:* Let $A(x) = 1 + \sum_{u=c}^d A_u x^u$, where $1 \leq c \leq d$, be a polynomial satisfying $A_c > 0$ and $A_u \geq 0$ for all $c < u \leq d$. Then

$$\lim_{\ell \to \infty} \frac{1}{\ell} \log \mathrm{Coeff} \left[ (A(x))^\ell, x^{\xi \ell} \right] = \log \left( \frac{A(z)}{z^\xi} \right) \tag{13}$$

where $z$ is the unique positive real solution to

$$\frac{A'(z)}{A(z)} \cdot z = \xi . \tag{14}$$

Applying this lemma by substituting $A(x) = A^{(t)}(x)$, $\ell = \gamma_t m$ and $\xi = \epsilon_t / \gamma_t$, we obtain that as $m \to \infty$

$$N_{c,t}^{(\gamma_t m)}(\epsilon_t m) = \mathrm{Coeff} \left[ \left( A^{(t)}(x) \right)^{\gamma_t m}, x^{\epsilon_t m} \right] \tag{15}$$

$$\to \exp \left\{ m \left( \gamma_t \log A^{(t)}(z_{0,t}) - \epsilon_t \log z_{0,t} \right) \right\} \tag{16}$$

where, for each $t \in I_c$, $z_{0,t}$ is the unique positive real solution to

$$\gamma_t \frac{\frac{\mathrm{d}A^{(t)}}{\mathrm{d}z}(z_{0,t})}{A^{(t)}(z_{0,t})} \cdot z_{0,t} = \epsilon_t . \tag{17}$$

The number of check-valid assignments of weight $\delta m$ satisfying the constraint $\boldsymbol{\epsilon}$ is obtained by multiplying the numbers of check-valid assignments of weight $\epsilon_t m$ over $\gamma_t m$ CNs of type $t$, for each $t \in I_c$,

$$N_c^{(\boldsymbol{\epsilon})}(\delta m) = \prod_{t \in I_c} N_{c,t}^{(\gamma_t m)}(\epsilon_t m) . \tag{18}$$

The number $N_c(\delta m)$ of check-valid assignments of weight $\delta m$ is then equal to the sum of $N_c^{(\boldsymbol{\epsilon})}(\delta m)$ over all admissible vectors $\boldsymbol{\epsilon}$; therefore by (16), as $m \to \infty$

$$N_c(\delta m) \to \sum_{\boldsymbol{\epsilon} \, : \, \sum_{t \in I_c} \epsilon_t = \delta} \exp \left\{ m W(\boldsymbol{\epsilon}) \right\} \tag{19}$$

where

$$W(\boldsymbol{\epsilon}) = \sum_{t \in I_c} \left( \gamma_t \log A^{(t)}(z_{0,t}) - \epsilon_t \log z_{0,t} \right) . \tag{20}$$

As $m \to \infty$, the asymptotic expression is dominated by the distribution $\boldsymbol{\epsilon}$ which maximizes the argument of the exponential function[2]. Therefore as $m \to \infty$

$$N_c(\delta m) \to \exp \left\{ m X \right\} \tag{21}$$

where

$$X = \max_{\boldsymbol{\epsilon}} W(\boldsymbol{\epsilon}) \tag{22}$$

and the maximization is subject to the constraint

$$V(\boldsymbol{\epsilon}) = \sum_{t \in I_c} \epsilon_t = \delta \tag{23}$$

together with $\epsilon_t \geq 0$ for each $t \in I_c$, and for every $t \in I_c$, $z_{0,t}$ is the unique positive real solution to (17). Note that for

each $t \in I_c$, (17) provides an implicit definition of $z_{0,t}$ as a function of $\epsilon_t$.

We solve this optimization problem using Lagrange multipliers, ignoring for the moment the inequality constraints. At the maximum, we must have

$$\frac{\partial W(\boldsymbol{\epsilon})}{\partial \epsilon_t} = \lambda \frac{\partial V(\boldsymbol{\epsilon})}{\partial \epsilon_t} \tag{24}$$

for all $t \in I_c$, where $\lambda$ is the Lagrange multiplier. This yields

$$\frac{\partial z_{0,t}}{\partial \epsilon_t} \left[ \gamma_t \frac{\frac{dA^{(t)}}{dz}(z_{0,t})}{A^{(t)}(z_{0,t})} - \frac{\epsilon_t}{z_{0,t}} \right] - \log z_{0,t} = \lambda . \tag{25}$$

The term in square brackets is equal to zero due to (17); therefore this simplifies to $\log z_{0,t} = -\lambda$ for all $t \in I_c$. We conclude that all of the $\{z_{0,t}\}$ are equal, and we may write

$$z_{0,t} = z_0 \quad \forall t \in I_c . \tag{26}$$

Making this substitution in (21) and using (23) we obtain

$$N_c(\delta m) \rightarrow \exp \left\{ m \left( \sum_{t \in I_c} \gamma_t \log A^{(t)}(z_0) - \delta \log z_0 \right) \right\} . \tag{27}$$

Summing (17) over $t \in I_c$ and using (23) and (26) implies that the value of $z_0$ in (27) is the unique positive real solution to (9) (here we have also used the fact that $n \int \rho = m \int \lambda$).

### B. Polynomial-System Solution for the Growth Rate

Consider the set of $\delta_t n$ VNs of a particular type $t \in I_v$, where $\delta_t$ is given by (1). Using generating functions, the number of variable-valid split assignments (over these VNs) of split weight $(\alpha_t n, \beta_t n)$ is given by

$$N_{v,t}^{(\delta_t n)}(\alpha_t n, \beta_t n) = \text{Coeff} \left[ \left( B^{(t)}(x,y) \right)^{\delta_t n}, x^{\alpha_t n} y^{\beta_t n} \right]$$

where $\text{Coeff}\,[p(x,y), x^c y^d]$ denotes the coefficient of $x^c y^d$ in the bivariate polynomial $p(x,y)$. We make use of the following result, which is a special case of [12, Corollary 16]:

*Lemma 2:* Let

$$B(x,y) = 1 + \sum_{u=1}^{k} \sum_{v=c}^{d} B_{u,v} x^u y^v$$

where $k \geq 1$ and $1 \leq c \leq d$, be a bivariate polynomial satisfying $B_{u,v} \geq 0$ for all $1 \leq u \leq k$, $c \leq v \leq d$. Then

$$\lim_{\ell \to \infty} \frac{1}{\ell} \log \text{Coeff} \left[ (B(x,y))^\ell, x^{\xi \ell} y^{\theta \ell} \right] = \log \left( \frac{B(x_0, y_0)}{x_0^\xi y_0^\theta} \right) \tag{28}$$

where $x_0$ and $y_0$ are the unique positive real solutions to the pair of simultaneous equations

$$\frac{\frac{\partial B}{\partial x}(x_0, y_0)}{B(x_0, y_0)} \cdot x_0 = \xi \tag{29}$$

and

$$\frac{\frac{\partial B}{\partial y}(x_0, y_0)}{B(x_0, y_0)} \cdot y_0 = \theta . \tag{30}$$

Applying this lemma by substituting $B(x,y) = B^{(t)}(x,y)$, $\ell = \delta_t n$, $\xi = \alpha_t/\delta_t$ and $\theta = \beta_t/\delta_t$, we obtain that as $n \to \infty$

$$N_{v,t}^{(\delta_t n)}(\alpha_t n, \beta_t n) = \text{Coeff} \left[ \left( B^{(t)}(x,y) \right)^{\delta_t n}, x^{\alpha_t n} y^{\beta_t n} \right]$$
$$\rightarrow \exp \left\{ n X_t^{(\delta_t)}(\alpha_t, \beta_t) \right\} \tag{31}$$

where

$$X_t^{(\delta_t)}(\alpha_t, \beta_t) = \delta_t \log B^{(t)}(x_{0,t}, y_{0,t}) - \alpha_t \log x_{0,t} - \beta_t \log y_{0,t} \tag{32}$$

and where $x_{0,t}$ and $y_{0,t}$ are the unique positive real solutions to the pair of simultaneous equations

$$\delta_t \frac{\frac{\partial B^{(t)}}{\partial x}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} \cdot x_{0,t} = \alpha_t \tag{33}$$

and

$$\delta_t \frac{\frac{\partial B^{(t)}}{\partial y}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} \cdot y_{0,t} = \beta_t . \tag{34}$$

Next, note that the expected number of D-GLDPC codewords of weight $\alpha n$ in the ensemble $\mathcal{M}_n$ is equal to the sum over $\beta$ of the expected numbers of split assignments of split weight $(\alpha n, \beta n)$ which are both check-valid and variable-valid, denoted $N_{\alpha n, \beta n}^{v,c}$:

$$\mathbb{E}_{\mathcal{M}_n}[N_{\alpha n}] = \sum_\beta \mathbb{E}_{\mathcal{M}_n}[N_{\alpha n, \beta n}^{v,c}] .$$

This may then be expressed as

$$\mathbb{E}_{\mathcal{M}_n}[N_{\alpha n}] = \sum_{\substack{\alpha_t \geq 0, t \in I_v \\ \sum_t \alpha_t = \alpha}} \sum_{\beta_t \geq 0, t \in I_v} P_{\text{c-valid}}(\beta n)$$
$$\times \prod_{t \in I_v} N_{v,t}^{(\delta_t n)}(\alpha_t n, \beta_t n) . \tag{35}$$

where $\beta = \sum_{t \in I_v} \beta_t$. Here $P_{\text{c-valid}}(\beta n)$ denotes the probability that a randomly chosen assignment of weight $\beta n$ is check-valid, and is given by

$$P_{\text{c-valid}}(\beta n) = N_c(\beta n) \Big/ \binom{E}{\beta n} .$$

Applying [12, eqn. (25)], we find that as $n \to \infty$

$$\binom{E}{\beta n} = \binom{n/\int \lambda}{\beta n} \rightarrow \exp \left\{ \frac{n}{\int \lambda} h \left( \beta \int \lambda \right) \right\} .$$

Combining this result with (27), we obtain that as $n \to \infty$

$$P_{\text{c-valid}}(\beta n) \rightarrow \exp \{ n Y(\beta) \}$$

where

$$Y(\beta) = \left( \frac{\int \rho}{\int \lambda} \right) \sum_{t \in I_c} \gamma_t \log \left( A^{(t)}(z_0) \right) - \beta \log z_0 - \frac{h(\beta \int \lambda)}{\int \lambda} .$$

Therefore, as $n \to \infty$

$$\mathbb{E}_{\mathcal{M}_n}[N_{\alpha n}] \rightarrow$$

$$\sum_{\substack{\alpha_t \geq 0, t \in I_v \\ \sum_t \alpha_t = \alpha}} \sum_{\beta_t, t \in I_v} \exp \left\{ n \left( \sum_{t \in I_v} X_t^{(\delta_t)}(\alpha_t, \beta_t) + Y(\beta) \right) \right\} \tag{36}$$

where

$$\beta \triangleq \sum_{t \in I_v} \beta_t \ . \tag{37}$$

Note that the sum in (36) is dominated asymptotically by the term which maximizes the argument of the exponential function. Thus, denoting the two vectors of independent variables by $\boldsymbol{\alpha} = (\alpha_t)_{t \in I_v}$ and $\boldsymbol{\beta} = (\beta_t)_{t \in I_v}$, we have

$$G(\alpha) = \max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} S(\boldsymbol{\alpha}, \boldsymbol{\beta}) \tag{38}$$

where

$$S(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{t \in I_v} X_t^{(\delta_t)}(\alpha_t, \beta_t) + Y(\boldsymbol{\beta}) \tag{39}$$

where $\beta$ is given by (37), and the maximization is subject to the constraint

$$R(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{t \in I_v} \alpha_t = \alpha \tag{40}$$

together with $\alpha_t \geq 0$ and appropriate inequality constraints on $\beta_t$ for each $t \in I_v$, and $\sum_t \alpha_t = \alpha$.

Note that (9) provides an implicit definition of $z_0$ as a function of $\boldsymbol{\beta}$. Similarly, for any $t \in I_v$, (33) and (34) provide implicit definitions of $x_{0,t}$ and $y_{0,t}$ as functions of the two variables $\alpha_t$ and $\beta_t$.

We solve the constrained optimization problem using Lagrange multipliers, ignoring for the moment the inequality constraints. At the maximum, we must have

$$\frac{\partial S(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_t} = \mu \frac{\partial R(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_t}$$

for all $t \in I_v$, where $\mu$ is the Lagrange multiplier. This yields

$$\frac{\partial x_{0,t}}{\partial \alpha_t} \left[ \delta_t \frac{\frac{\partial B^{(t)}}{\partial x}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} - \frac{\alpha_t}{x_{0,t}} \right] - \log x_{0,t}$$
$$+ \frac{\partial y_{0,t}}{\partial \alpha_t} \left[ \delta_t \frac{\frac{\partial B^{(t)}}{\partial y}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} - \frac{\beta_t}{y_{0,t}} \right] = \mu \ .$$

The terms in square brackets are zero due to (33) and (34) respectively; therefore this simplifies to $\log x_{0,t} = -\mu$ for all $t \in I_v$. We conclude that all of the $\{x_{0,t}\}$ are equal, and we may write

$$x_{0,t} = x_0 \quad \forall t \in I_v \ . \tag{41}$$

At the maximum, we must also have

$$\frac{\partial S(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_t} = \mu \frac{\partial R(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_t}$$

for all $t \in I_v$. This yields

$$\frac{\partial x_{0,t}}{\partial \alpha_t} \left[ \delta_t \frac{\frac{\partial B^{(t)}}{\partial x}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} - \frac{\alpha_t}{x_{0,t}} \right] - \log y_{0,t} - \log z_0$$
$$+ \frac{\partial y_{0,t}}{\partial \beta_t} \left[ \delta_t \frac{\frac{\partial B^{(t)}}{\partial y}(x_{0,t}, y_{0,t})}{B^{(t)}(x_{0,t}, y_{0,t})} - \frac{\beta_t}{y_{0,t}} \right] - \log \left( \frac{1 - \beta \int \lambda}{\beta \int \lambda} \right)$$
$$+ \frac{\partial z_0}{\partial \beta_t} \left[ \left( \frac{\int \rho}{\int \lambda} \right) \sum_{s \in I_c} \gamma_s \frac{\frac{dA^{(s)}}{dz}(z_0)}{A^{(s)}(z_0)} - \frac{\beta}{z_0} \right] = 0 \ . \tag{42}$$

The terms in square brackets are zero due to (33), (34) and (9) respectively; therefore this simplifies to

$$z_0 y_{0,t} \left( \frac{1 - \beta \int \lambda}{\beta \int \lambda} \right) = 1 \quad \forall t \in I_v \ . \tag{43}$$

We conclude that all of the $\{y_{0,t}\}$ are equal, and we may write

$$y_{0,t} = y_0 \quad \forall t \in I_c \ . \tag{44}$$

Rearranging (43) we obtain (12). Also, summing (33) over $t \in I_v$ and using (40) and (41) yields (10). Similarly, summing (34) over $t \in I_v$ and using (37) and (44) yields (11). Substituting back into (39) and using (41), (44), (40) and (37) yields

$$G(\alpha) = \sum_{t \in I_v} \delta_t \log B^{(t)}(x_0, y_0) - \alpha \log x_0 - \beta \log y_0$$
$$+ \left( \frac{\int \rho}{\int \lambda} \right) \sum_{s \in I_c} \gamma_s \log A^{(s)}(z_0) - \beta \log z_0 - \frac{h(\beta \int \lambda)}{\int \lambda} \tag{45}$$

where $x_0$, $y_0$, $z_0$ and $\beta$ are the unique positive real solutions to the $4 \times 4$ system of equations (9), (10), (11) and (12). Finally, (12) leads to the observation that

$$-\beta \log z_0 - \beta \log y_0 - \frac{h(\beta \int \lambda)}{\int \lambda} = \frac{\log \left( 1 - \beta \int \lambda \right)}{\int \lambda}$$

which, when substituted in (45), leads to (8).

## V. EXAMPLES

In this section the growth rates of two example D-GLDPC ensembles of design rate $R = 1/2$ are evaluated using the polynomial solution of Theorem 1. We use Hamming $(7,4)$ codes as generalized CNs and SPC codes as generalized VNs. Three representations of SPC VNs are considered, namely, the cyclic (C), the systematic (S) and the antisystematic (A) representations [3].

Ensemble 1 is characterized by two CN types and two VN types. Specifically, we have $I_c = \{1, 2\}$, where $1 \in I_c$ denotes a $(7, 4)$ Hamming CN type and $2 \in I_c$ denotes a length-7 single parity check (SPC) CN type, and $I_v = \{1, 2\}$, where $1 \in I_v$ denotes a repetition-2 VN type and $2 \in I_v$ denotes a length-7 SPC CN type in cyclic form. Ensemble 2 is characterized by two CN types and four VN types. Specifically, we have $I_c = \{1, 2\}$, where $1 \in I_c$ denotes a $(7, 4)$ Hamming CN type and $2 \in I_c$ denotes a SPC-7 CN type, and $I_v = \{1, 2, 3, 4\}$, where $1 \in I_v$ denotes a repetition-2 VN type, $2 \in I_v$ denotes a length-7 SPC CN type in cyclic form, $3 \in I_v$ denotes a length-7 SPC CN type in antisystematic form, and $4 \in I_v$ denotes a length-7 SPC CN type in systematic form. The edge-perspective type distributions of the two ensembles are summarized in Table I.

Both Ensemble 1 and Ensemble 2 have been obtained by performing a decoding threshold optimization with differential evolution (DE) [20]. Ensemble 1 has been obtained by only

---

[3]The $(k \times (k+1))$ generator matrix of a SPC code in A form is obtained from the generator matrix in S form by complementing each bit in the first $k$ columns. Note that a $(k \times (k+1))$ generator matrix in A form represents a SPC code if and only if the code length $q = k+1$ is odd. For even $k+1$ we obtain a $d_{\min} = 1$ code with one codeword of weight 1.

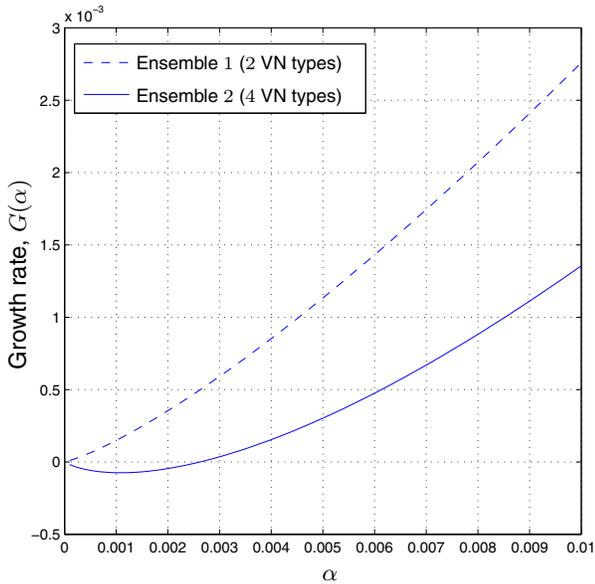| Ensemble 1 | | | |
|---|---|---|---|
| Variable nodes | | Check nodes | |
| 1:repetition$-2$ | $\lambda_1 = 0.055646$ | 1:Hamming$(7,4)$ | $\rho_1 = 0.965221$ |
| 2:SPC$-7$ (C) | $\lambda_2 = 0.944354$ | 2:SPC$-7$ | $\rho_2 = 0.034779$ |
| Ensemble 2 | | | |
| Variable nodes | | Check nodes | |
| 1:repetition$-2$ | $\lambda_1 = 0.022647$ | 1:Hamming$(7,4)$ | $\rho_1 = 0.965221$ |
| 2:SPC$-7$ (C) | $\lambda_2 = 0.100000$ | 2:SPC$-7$ | $\rho_2 = 0.034779$ |
| 3:SPC$-7$ (A) | $\lambda_2 = 0.539920$ | | |
| 4:SPC$-7$ (S) | $\lambda_2 = 0.337432$ | | |



Fig. 1. Growth rates of the two example ensembles described in Section V. Ensemble 1 is asymptotically bad, while Ensemble 2 is asymptotically good with an ensemble relative minimum distance of $\alpha^* = 2.625 \times 10^{-3}$.

imposing the node type and $R = 1/2$ constraint. In this case we have $C \cdot V = 1.19 > 1$, so the ensemble is *asymptotically bad* ($\alpha^* = 0$). Ensemble 2 has been obtained by imposing the node type and $R = 1/2$ constraint, together with the constraints $C \cdot V \leq 0.5$ and $\lambda_2 \geq 0.1$. Since in this case we have $C \cdot V = 0.5 < 1$, the ensemble is *asymptotically good* ($\alpha^* > 0$). The expected asymptotically bad or good behavior of the two ensembles is reflected in the growth rate curves shown in Fig. 1. Using a standard numerical solver, it took only 5.1 s and 6.7 s to evaluate 100 points on the Ensemble 1 curve and on the Ensemble 2 curve, respectively. The relative minimum distance of Ensemble 2 is $\alpha^* = 2.625 \times 10^{-3}$.

## VI. CONCLUSION

A general expression for the asymptotic growth rate of the weight distribution of irregular D-GLDPC ensembles has been presented. Evaluation of the expression requires solution of a $4 \times 4$ polynomial system, irrespective of the number of VN and CN types in the ensemble. Simulation results were presented for two example optimized irregular D-GLDPC code ensembles.

## REFERENCES

[1] R. Gallager, *Low-Density Parity-Check Codes*. Cambridge, Massachusetts: M.I.T. Press, 1963.
[2] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533–547, Sept. 1981.
[3] G. Liva, W. E. Ryan, and M. Chiani, "Quasi-cyclic generalized LDPC codes with low error floors," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 49–57, Jan. 2008.
[4] Y. Wang and M. Fossorier, "Doubly generalized low-density parity-check codes," in *Proc. of 2006 IEEE Int. Symp. on Information Theory*, Seattle, WA, USA, July 2006, pp. 669–673.
[5] N. Miladinovic and M. Fossorier, "Generalized LDPC codes and generalized stopping sets," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 201–212, Feb. 2008.
[6] E. Paolini, M. Fossorier, and M. Chiani, "Doubly-generalized LDPC codes: Stability bound over the BEC," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1027–1046, Mar. 2009.
[7] M. Lentmaier and K. Zigangirov, "On generalized low-density parity-check codes based on Hamming component codes," *IEEE Commun. Lett.*, vol. 3, no. 8, pp. 248–250, Aug. 1999.
[8] J. Boutros, O. Pothier, and G. Zemor, "Generalized low density (Tanner) codes," in *Proc. of 1999 IEEE Int. Conf. on Communications*, vol. 1, Vancouver, Canada, June 1999, pp. 441–445.
[9] S. Litsyn and V. Shevelev, "On ensembles of low-density parity-check codes: asymptotic distance distributions," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 887–908, Apr. 2002.
[10] D. Burshtein and G. Miller, "Asymptotic enumeration methods for analyzing LDPC codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1115–1131, June 2004.
[11] A. Orlitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 929–953, Mar. 2005.
[12] C. Di, R. Urbanke, and T. Richardson, "Weight distribution of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4839–4855, Nov. 2006.
[13] D. Divsalar, "Ensemble weight enumerators for protograph LDPC codes," in *2006 IEEE Int. Symp. on Information Theory*, Seattle, WA, USA, July 2006, pp. 1554–1558.
[14] T. Awano, K. Kasai, T. Shibuya, and K. Sakaniwa, "Three-edge type LDPC code ensembles with exponentially few codewords with linear small weight," in *2008 Int. Symp. on Information Theory and its Applications*, Auckland, New Zealand, Dec. 2008, pp. 731–735.
[15] S. Abu-Surra, W. E. Ryan, and D. Divsalar, "Ensemble enumerators for protograph-based generalized LDPC codes," in *Proc. of 2007 IEEE Global Communications Conf.*, Washington, DC, USA, Nov. 2007, pp. 1492–1497.
[16] Y. Wang, C.-L. Wang, and M. Fossorier, "Ensemble weight enumerators for protograph-based doubly generalized LDPC codes," in *Proc. of 2008 IEEE Int. Symp. on Information Theory*, Toronto, Canada, July 2008, pp. 1168–1172.
[17] M. F. Flanagan, E. Paolini, M. Chiani, and M. Fossorier, "On the growth rate of the weight distribution of irregular doubly-generalized LDPC codes," in *in Proc. of the 46-th Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, USA, Sept. 2008, pp. 922–929.
[18] E. Paolini, M. F. Flanagan, M. Chiani, and M. Fossorier, "On a class of doubly-generalized LDPC codes with single parity-check variable nodes," in *Proc. of 2009 IEEE Int. Symp. on Information Theory*, Seoul, Korea, July 2009, pp. 1983–1987.
[19] M. Flanagan, E. Paolini, M. Chiani, and M. Fossorier, "On the growth rate of the weight distribution of irregular doubly-generalized LDPC codes," *IEEE Trans. Inf. Theory*, 2009, submitted.
[20] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, 1st ed. Berlin, Germany: Springer-Verlag, 2005.