Exact Thresholds and Optimal Codes for the Binary Symmetric Channel and Gallager's Decoding Algorithm A

Louay Bazzi, Tom Richardson and Rüdiger Urbanke
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974

August 19, 1999

Abstract

We show that for the case of a binary symmetric channel and Gallager's decoding algorithm A the threshold can, in many cases, be determined analytically. More precisely, we show that the threshold is always upper bounded by the minimum of $\frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)}$ and the smallest positive real root τ of a specific polynomial $p(x)$ and we observe that for most cases this bound is tight, i.e., it determines the threshold exactly. We further present optimal codes for a large range of rates. In the case of rate one-half codes, for example, the threshold x_0^* of the optimal code is given by $x_0^* \sim 0.0513663$. Finally, we outline how thresholds of more complicated decoders might be determined analytically.

1 Introduction

In his thesis (1962) Gallager introduced the concept of low density parity check codes (LDPCCs) together with various iterative decoding algorithms. For an introduction into the basic concepts of LDPCCs we refer the reader to [1, 2, 3, 4, 5]. The iterative decoding algorithms are best interpreted as *message passing* algorithms executed on a bipartite graph which represents the code. Messages are passed along the edges of the graph between *variable* nodes on the left hand side and *check* nodes on the right. The variable nodes are in one-to-one correspondence with the bits of the codeword and the check nodes are in one-to-one correspondence with the set of linear constraints on the bits of the code which determine the code. Decoding generally begins with *variable* nodes sending to their neighboring check nodes their *received* message. This message corresponds with the observation of the associated bit after passing through a communications channel. Decoding continues with rounds of messages sent back and forth from check nodes to variable nodes where new messages are computed by each node as functions of messages previously sent to them. Generally speaking, the messages represent estimates of the bit associated to the variable node terminal to the edge carrying the message.

The best, but also most complex, iterative decoding algorithm proposed in [1] is an instance of belief propagation [6], a message passing algorithm in which the messages represent conditional probabilities. The simplest decoding algorithm, which we will call Gallager's decoding algorithm A, allows as messages only elements of the set $\{\pm 1\}$ indicating the *sign* of a bit. (Here we assume that the communications channel is the binary symmetric channel.) More precisely, at a *check* node the outgoing message along an edge e is simply the product of all incoming messages excluding the message incoming along edge e. At a *variable* node the outgoing message is equal to the received message except if all incoming messages agree in which case this message is forwarded. Although this simple decoding algorithm can clearly not match the performance of belief propagation it is nevertheless of interest because of its extremely low complexity.

In Gallager's original description every variable node participates in the same number j of checks and every check involves the same number k of variable nodes. This gives rise to what is now commonly known as a *regular* LDPCC. In [3, 2] it was suggested to let variables participate in variable number of checks and to allow varying sizes of check sets. Such codes are called *irregular* LDPCCs. Hence, in general LDPCCs are characterized by a *degree sequences* pair $(\lambda(x), \rho(x))$ where $\lambda(x)$ and $\rho(x)$ are functions of the form $\lambda(x) := \sum_{i=2}^{\infty} \lambda_i x^{i-1}$, $\rho(x) := \sum_{i=2}^{\infty} \rho_i x^{i-1}$ and where λ_i (ρ_i) denotes the fraction of ones in the parity check matrix of the code which are in columns (rows) of weight i.

Example 1. [(3, 6) code] For the (3, 6)-regular LDPCC we have $\lambda(x) := x^2$ and $\rho(x) := x^5$.

An analysis of message passing algorithms is possible under an *independence assumption*. One analyzes locally the distribution of outgoing messages assuming all incoming messages are independent. If, in a finite graph, one *unvolves* the graph relative to a particular message then one obtains a *tree* structure representing the history of the message. For finite graphs this tree will have within it repetitions, i.e., variable nodes, check nodes, and edges may appear more than once. The independence assumption is equivalent to assuming that there are no repetitions in the tree. If one considers a finite number of iterations of the message passing algorithm then asymptotically, as the length of the code tends to infinity, most messages will satisfy the independence assumption.

For the binary symmetric channel and various regular codes Gallager numerically determined the largest cross-over probability such that, under the independence assumption, the expected number messages in error converges to zero as the number of iterations grows to infinity. It was shown in [4] that this number, which is called the *threshold*, determines the asymptotic (in the codeword length) behavior of the ensemble of codes: roughly speaking, for a code chosen randomly from the ensemble, with high probability, decoding will be successful if transmission takes place below this threshold, and the error probability will stay above a fixed constant if transmission takes place above this threshold.

Although this threshold can be easily determined numerically it is nevertheless pleasing and potentially insightful to derive analytic expressions. It is the aim of this correspondence to show that for the binary symmetric channel and Gallager's decoding algorithm A the threshold can be given explicitly in many cases. Moreover, for a wide range of rates the optimal codes can be determined. The techniques used in determining the thresholds analytically are best illustrated by means of the even simpler example of the binary erasure channel (BEC). Let x_0 be the initial number of erasures. Then the expected number of erasures fulfills the recursion [3]

$$x_l = x_0\lambda(1-\rho(1-x_{l-1})). \tag{1}$$

We are interested in the maximal value of x_0 such that x_l converges to zero as l tends to infinity. It is easily seen that this value of x_0, call it x_0^*, is equal to the supremum of all values x_0 such that the equation

$$x = x_0\lambda(1-\rho(1-x)) \tag{2}$$

does not have a positive solution x with $x \le x_0$. Solving equation (2) for x_0 we get

$$x_0(x) = \frac{x}{\lambda(1-\rho(1-x))}. \tag{3}$$

In words, for a given positive real number x there is a unique value of x_0 such that x fulfills the fixed point equation (2). By the remarks above, if $x \le x_0$ then the threshold is upper bounded by x_0. This fact is used in the following lemma to determine the threshold of regular codes for the BEC.

Lemma 1. [Thresholds for the BEC and Regular Codes] Assume we are given a (j,k)-regular LDPCC. Let σ be the unique positive real root of the polynomial $p(x) = ((j-1)(k-1)-1)x^{k-2} - \sum_{i=0}^{k-3} x^i$. Then the threshold $x_0^*(j,k)$ is equal to $x_0^*(j,k) = \frac{1-\sigma}{(1-\sigma^{k-1})^{j-1}}$.

Proof. By the remarks above, the threshold is given by $\min\{g(x) : g(x) \ge x\}$, where $g(x) := \frac{x}{\lambda(1-\rho(1-x))} = \frac{x}{(1-(1-x)^{k-1})^{j-1}}$. First note that $g(x) \ge x$ with equality only at $x = 1$. It follows that the threshold is given by the minimum of $g(x)$ over the range $x \in [0,1]$. In the sequel it will be slightly more convenient to consider $g(1-x) = \frac{1-x}{(1-x^{k-1})^{j-1}}$. Taking the derivative of $g(1-x)$ with respect to x and setting the result to zero we get the polynomial equation

$$((k-1)(j-1)-1)x^{k-1} - (k-1)(j-1)x^{k-2} + 1 = 0.$$

By Descartes' rule of signs[1] this equation has at most two positive real roots. Clearly, one such root is at $x = 1$. Hence, dividing by $x - 1$ we get the simplified polynomial equation

$$((j-1)(k-1)-1)x^{k-2} - \sum_{i=0}^{k-3} x^i = 0,$$

[1]Descartes' rule of signs bounds the number of positive real zeros of a polynomial $f(x)$ in one variable. If

$$f(x) = \sum_{j=1}^{r} c_j x^{m_j} \;,$$

with $0 \le m_1 < m_2 < \ldots < m_r$ and with all coefficients $c_j \ne 0$, then the number of positive real zeros of f is upper bounded by the number of sign changes $N^+(f)$ between consecutive coefficients c_j when taken in order of increasing j, see [7], [8, Chapter 6].

which by the above remark has at most a single real positive root. It is easy to see that $g(x)$ has a pole at $x = 0$ and is equal to 1 at $x = 1$. Hence, $g(x)$ does not take on its minimum at the boundary. It follows that it must take on its minimum in the interior, hence, at least one and therefore exactly one solution to the above polynomial equation must exist. $\square$

In the following examples the threshold for some standard regular codes are determined.

Example 2. [(3, 6) code] Let σ be given by

$$\sigma = \frac{1}{36} + \frac{\sqrt{\frac{25}{324} - a + b}}{2} + \frac{\sqrt{\frac{25}{162} + a - b + \frac{685}{2916\sqrt{\frac{25}{324} - a + b}}}}{2},$$

where $a = \frac{22}{27}5^{\frac{2}{3}}\left(\frac{2}{-85+3\sqrt{24465}}\right)^{\frac{1}{3}}$ and $b = \frac{1}{27}\left(\frac{5}{2}\left(-85 + 3\sqrt{24465}\right)\right)^{\frac{1}{3}}$. Then $x_0^*(3,6) := \frac{1-\sigma}{(1-\sigma^5)^2} \sim 0.42944$.

Example 3. [(3, 5) code] $x_0^*(3,5) = \frac{1+\frac{-1-(694-42\sqrt{267})^{\frac{1}{3}}-(694+42\sqrt{267})^{\frac{1}{3}}}{21}}{\left(1-\frac{\left(1+(694-42\sqrt{267})^{\frac{1}{3}}+(694+42\sqrt{267})^{\frac{1}{3}}\right)^4}{194481}\right)^2} \sim 0.540605.$

Example 4. [(4, 6) code] Let σ be given by

$$\sigma = \frac{1}{56} + \frac{1}{2}\sqrt{\frac{115}{2352} - a + b} + \frac{1}{2}\sqrt{\frac{115}{1176} + a - b + \frac{1625}{10976\frac{115}{2352} - a + b}},$$

where $a = \frac{175^{2/3}}{21(-65+3\sqrt{22305})^{1/3}}$ and $b = \frac{1}{42}(5(-65 + 3\sqrt{22305}))^{1/3}$. Then $x_0^*(4,6) := \frac{1-\sigma}{(1-\sigma^5)^3} \sim 0.506741$.

Example 5. [(3, 4) code] $x_0^*(3,4) = \frac{3125}{3672+252\sqrt{21}} \sim 0.647426$.

It was indicated in [4] that simple extensions to Gallager's decoding algorithm A involving message alphabets of size only three or four can reach large fractions of the capacity at very low complexity. For example, a simple *decoder with erasures* in which messages take on values in $\{-1, 0, 1\}$, with 0 indicating an erasure, yields a threshold for the $(3,6)$-regular code of roughly 0.07. This should to be compared to the threshold under belief propagation 0.084 and also to Shannon capacity of 0.11. An optimization with respect to the code would almost certainly increase this threshold even further, as would an extension of the message alphabet to size four. Such decoders are prime candidates for practical implementations, yielding large fractions of the capacity at extremely low complexities. Section 4 contains a conjecture concerning the exact threshold of the $(3,6)$-regular code for this decoder with erasures. It remains a challenge for future research to see if the thresholds for more complex decoding schemes can still be determined analytically and if the corresponding optimal codes can be found.

2 Exact Thresholds

We now show that for the case of a binary symmetric channel and Gallager's decoding algorithm A the threshold can, in most cases, be determined analytically.

Let x_0 be the expected number of initial errors, i.e., x_0 equals the cross-over probability of the binary symmetric channel. It was shown by Gallager [1, p. 48] that the expected number of errors (under the independence assumption) in the l-th iteration is given by the recursion

$$x_l = x_0 - x_0 p^+(x_{l-1}) + (1 - x_0)p^-(x_{l-1}), \tag{4}$$

where

$$\begin{aligned} p^+(x) &:= \lambda\left(\frac{1+\rho(1-2x)}{2}\right), \\ p^-(x) &:= \lambda\left(\frac{1-\rho(1-2x)}{2}\right), \end{aligned}$$

and where $(\lambda(x), \rho(x))$ is the degree sequence pair.

Definition 1. The threshold $\tilde{x}_0^*$ is the supremum of all x_0 in $\mathbb{R}^+$ such that $x_l(x_0)$ as defined in (4) converges to zero as l tends to infinity.

Let us define

$$f(x, y) := x - xp^+(y) + (1 - x)p^-(y).$$

For small y we may write

$$f(x, y) = x\left[\lambda'(1)\rho'(1) - \lambda_2\rho'(1)\right] y + \lambda_2\rho'(1)y + O(y^2),$$

where $\lambda'(x)$ and $\rho'(x)$ denote derivatives of $\lambda(x)$ and $\rho(x)$, respectively. Since $(\lambda'(1) - \lambda_2)\rho'(1) \geq 0$ we see that unless $\lambda_2\rho'(1) \leq 1$ we will have $\tilde{x}_0^* = 0$. If $\lambda_2\rho'(1) = 1$ then, since $\rho'(1) > 1$, we have $(\lambda'(1) - \lambda_2)\rho'(1) > 0$ and again we will have $\tilde{x}_0^* = 0$. Thus, a necessary condition for $\tilde{x}_0^* > 0$ is

$$\lambda_2\rho'(1) < 1, \tag{5}$$

and we shall assume that this holds henceforth.

Lemma 2. *In the range $0 \leq x \leq 1$ and $0 \leq y \leq \frac{1}{2}$, f is an increasing function in both its arguments. Furthermore, if $x < \frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)}$ then for small enough y, $f(x, y) < y$.*

Proof. We start by showing that $f(x, y)$ is an increasing function in x. Note that $p^+(y) + p^-(y) = \lambda(q(y)) + \lambda(1 - q(y))$, where $0 \leq q(y) = \frac{1+\rho(1-2y)}{2} \leq 1$, for $0 \leq y \leq \frac{1}{2}$. Since λ has a Taylor series expansion with non-negative coefficients and $\lambda(1) = 1$, it follows that $1 - p^+(y) - p^-(y) \geq 0$. The claim now follows by observing that $f(x, y) := x(1 - p^+(y) - p^-(y)) + p^-(y)$.

Taking the derivative of $f(x, y)$ with respect to y we get

$$\frac{\partial}{\partial y} f(x, y) = \left[x\lambda'\left(\frac{1+\rho(1-2y)}{2}\right) + (1 - x)\lambda'\left(\frac{1-\rho(1-2y)}{2}\right)\right]\rho'(1 - 2y).$$

If $0 \le x \le 1$ and $0 \le y \le \frac{1}{2}$ then this expression is positive which proves that f is an increasing function in its second argument.

If $x = \frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)} - \epsilon$, $\epsilon > 0$, then

$$\begin{aligned} f(x,y) &= x\left[\lambda'(1)\rho'(1) - \lambda_2\rho'(1)\right] y + \lambda_2\rho'(1)y + O(y^2) \\ &< \left(1 - \epsilon\left[\lambda'(1) - \lambda_2\right]\rho'(1)\right) y + O(y^2), \end{aligned}$$

which proves the last statement. □

Since f is an increasing function in both its arguments it follows by finite induction that for any $x_0 < \tilde{x}_0^*$ the sequence $x_l(x_0)$ will converge to zero. Moreover, if $x_l(x_0)$ converges to zero then $x_l(x_0)$ must be strictly decreasing. In the sequel it will be more convenient to work with the following alternative description of the threshold.

Definition 2. The threshold x_0^* is the supremum of all x_0 in $\mathbb{R}^+$ such that

$$x = x_0 - x_0 p^+(x) + (1 - x_0)p^-(x). \tag{6}$$

does not have a strictly positive solution x with $x \le x_0$.

Note that since $p^+(\frac{1}{2}) = p^-(\frac{1}{2}) = \lambda(\frac{1}{2})$ we have $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$, hence $x_0^* \le \frac{1}{2}$.

Lemma 3. Assuming (5), we have $x_0^* = \tilde{x}_0^*$.

Proof. We first show that $\tilde{x}_0^* \ge x_0^*$. By definition of x_0^* we have $f(x,x) - x \ne 0$ for $x \in (0, x_0^*)$. By Lemma 2, $f(x,x) - x < 0$ for x sufficiently small and, since $f(x,x) - x$ is continuous in x, we have $x_1(x) = f(x,x) < x$ for every $x \in (0, x_0^*)$. Let $x_0 < x_0^*$. Since $x_1(x_0) < x_0 \le \frac{1}{2}$ it follows from Lemma 2 that $x_l(x_0)$ is non-increasing and therefore converges to some value x_∞ which satisfies $f(x_0, x_\infty) = x_\infty$. Since, by assumption, there are no solutions to $f(x_0, x) = x$ for $x \in (0, x_0]$ we conclude that $x_\infty = 0$. Thus, we now have $\tilde{x}_0^* \ge x_0^*$.

We now show that $\tilde{x}_0^* \le x_0^*$. By the definition of x_0^* it follows that we can find $x_0 \in [x_0^*, \frac{1}{2}]$ arbitrarily close to x_0^* such that $f(x_0, x) = x$ for some $x \in (0, x_0]$. I.e., x fulfills (6) for x_0. We will show that $\tilde{x}_0^* \le x_0$, completing the proof. For any $y \in [x, \frac{1}{2}]$ we have by Lemma 2 that $x = f(x_0, x) \le f(x_0, y) \le f(x_0, \frac{1}{2}) \le \frac{1}{2}$. Since $x_0 \in [x, \frac{1}{2}]$ we have $x_l(x_0) \in [x, \frac{1}{2}]$ for all l and it follows that $x_0 \ge \tilde{x}_0^*$. □

Rather than solving equation (6) for x for a given x_0 we note that, generically, for a given x there is a unique solution for x_0 given by

$$x_0(x) = \frac{x - p^-(x)}{1 - p^+(x) - p^-(x)} := g(x).$$

Note that $p^+(x) - 1 + x > 0$ for $x \in (0, 1/2)$ since $p^+(x) - 1 - x$ is convex (differentiate twice) and evaluates to 0 and 1/2 at $x = 0$ and $x = 1/2$ respectively. Thus, if $1 - p^+(x) - p^-(x) = 0$ for $x \in (0, 1/2)$, then we have $x - p^-(x) \ne 0$ and we may uniquely define $x_0(x) \in \{+\infty, -\infty\}$.

For $x = 0$, continuity dictates that $g(x)$ is defined in terms of right limits and we get $g(0) := \frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)}$. According to Definition 2 we have

$$x_0^* := \min_{0 \le x}\{g(x) : g(x) \ge x\}. \tag{7}$$

This immediately gives rise to the following upper bound on the threshold.

d_l	d_r	Rate	x_0^*
3	6	0.5	$\tau_{(3,6)}$
4	8	0.5	$\frac{1}{21}$
5	10	0.5	$\frac{1}{36}$
3	5	0.4	$\tau_{(3,5)}$
4	6	0.333	$\frac{1}{15}$
3	4	0.25	$\tau_{(3,4)}$

Table 1: Thresholds for the binary symmetric channel and Gallager's decoding algorithm A for some standard regular codes.

Lemma 4. Let τ denote the smallest positive real root of the polynomial $p(x) := xp^+(x)+(x-1)p^-(x)$ and assume (5) holds. Then $x_0^* \leq X_0^* := \min\{\frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)}, \tau\}$.

Proof. Evaluation of $g(x)$ at any point $x \geq 0$ where $g(x) \geq x$ gives an upper bound for x_0^*. In particular we evaluate $g(x)$ at $x = 0$ and at the smallest positive solution of the equation $x = g(x)$. Taking the minimum of these two values gives our bound. □

We note that, although one can construct counterexamples, for most codes the minimum is taken on at one of the two endpoints. Observe that the bound which stems from the evaluation at $x = 0$ can be seen as a *stability* condition, see [5], whereas the bound which stems from the evaluation at the right end point ensures that the iterative process gets started.

Table 1 summarizes thresholds of some standard regular codes. In all these cases the minimum in (7) is taken on either at the left or at the right endpoint. For the $(4,8)$, $(5,10)$ and $(4,6)$ code the minimum is taken on at the left endpoint so that the threshold is equal to $g(0)$. For the remaining cases the minimum is taken at the right endpoint so that the threshold is given in terms of the smallest positive real root of the polynomial $p(x)$. The latter cases are treated in some more detail in the following examples.

Example 6. [(3, 6) code] For the $(3,6)$ code we have

$$p(x) := (-1+x)x(-1+2x)(1-32x+192x^2-640x^3+1376x^4-1984x^5+1856x^6-1024x^7+256x^8).$$

Note that

$$p(\frac{1-x}{2})\frac{4}{(1-x)x(1+x)} = -1-x^2+x^4+x^6+x^8 = r(x^2),$$

where $r(x) := -1-x+x^2+x^3+x^4$. Applying Descartes' rule of signs [7], we observe that $r(x)$ has a single positive root which we denote by σ. Then

$$\tau_{(3,6)} = \frac{1-\sqrt{\sigma}}{2}.$$

The polynomial can be solved by radicals and the solution is given by

$$\sigma := -\frac{1}{4} + \frac{\sqrt{-\frac{5}{12}-\omega}}{2} + \frac{\sqrt{-\frac{5}{6}+\frac{11}{4\sqrt{-\frac{5}{12}-\omega}}+\omega}}{2},$$

where

$$\omega := \frac{8}{3}\left(\frac{2}{83+3\sqrt{993}}\right)^{\frac{1}{3}} - \frac{1}{3}\left(\frac{83+3\sqrt{993}}{2}\right)^{\frac{1}{3}}.$$

Numerically, the first 10 digits of $\tau_{(3,6)}$ are $\tau_{(3,6)} = 0.0394636562$.

Example 7. [(3, 5) code] In this case we have

$$p(x) = x\,(-1+2\,x)\,\left(-1+22\,x-108\,x^2+280\,x^3-448\,x^4+448\,x^5-256\,x^6+64\,x^7\right).$$

This can be slightly simplified by noting that

$$p\left(\frac{1-x}{2}\right) = \frac{x(1-x)}{4}(x^7+x^6+x^5+x^4+x^3-x^2-x^1-1).$$

By Descartes rule of signs the last factor has at most one positive real root, and it clearly has one root in $(0,1)$. In this case it might not be possible to express this root in terms of radicals but $\tau_{(3,5)}$ can of course be determined to any desired degree by a standard root extraction algorithm. The first 10 digits of $\tau_{(3,5)}$ are given by $\tau_{(3,5)} = 0.0611860546$.

Example 8. [(3, 4) code] For the $(3,4)$ code we have

$$p(x) = (-1+x)\,x\,(-1+2\,x)\,\left(1-12\,x+28\,x^2-32\,x^3+16\,x^4\right).$$

The smallest positive real root is a root of the last factor and is given by

$$\tau_{(3,4)} = \frac{2-\sqrt{2\,\left(-1+\sqrt{5}\right)}}{4}.$$

3 Optimal Codes

Given the ease with which thresholds can be determined, one might wonder whether optimal codes for the given decoder can be found. We will show in this section that for a wide range of rates this is indeed the case. The proof has two parts. First, we show how to find optimal *right-concentrated* degree sequence pairs, i.e., degree sequence pairs (λ,ρ) for which ρ has at most two non-zero degrees and for which these non-zero degrees are consecutive. Second, we prove the optimality of such right-concentrated degree sequence pairs for a wide range of rates.

We will make repeated use of the following basic fact of linear programming: Given a primal and a dual problem and a feasible solution for each which have the same value, these feasible solutions are optimal [9].

We start by showing that for a given right degree sequence ρ the optimal left degree sequence λ has bounded degree. We will use the following notation. For any function $\lambda(x) = \sum_{i=2}^{\infty}\lambda_i x^{i-1}, \lambda_i \geq 0$ let $\int\lambda$ abbreviate $\int_0^1 \lambda(x)\,dx = \sum_{i=2}^{\infty}\lambda_i i$. If (λ,ρ) is a degree sequence pair then $(\int\lambda\,)^{-1}, (\int\rho\,)^{-1}$ are the average degrees of variable nodes and check nodes respectively.

Theorem 1. [Bounded Left Degree] Let (λ,ρ) be a given degree sequence pair of rate $r(\lambda,\rho)$ and threshold $x_0^*(\lambda,\rho)$. Then there exists a degree sequence pair $(\tilde{\lambda},\rho)$ such that

i) The largest index of any non-zero degree of $\tilde{\lambda}$ is at most $i_{\max} = \lceil \frac{\log \frac{1-x_0^*}{x_0^*}}{\log \frac{1+\rho(1-2x_0^*)}{1-\rho(1-2x_0^*)}} \rceil + 2$.

ii) $r(\tilde{\lambda}, \rho) \geq r(\lambda, \rho)$.

iii) $x_0^*(\tilde{\lambda}, \rho) \geq x_0^*(\lambda, \rho)$.

Proof: For $i > 1$, let

$$h_i(x) := x_0^* \left(\frac{1 + \rho(1-2x)}{2} \right)^{i-1} - (1 - x_0^*) \left(\frac{1 - \rho(1-2x)}{2} \right)^{i-1}.$$

Let $x \in (0, x_0^*)$. Then by assumption we have $x > x_0^* - x_0^* \lambda \left(\frac{1+\rho(1-2x)}{2} \right) + (1 - x_0^*) \lambda \left(\frac{1-\rho(1-2x)}{2} \right)$, which in terms of the h_i's can be expressed as $\sum_{i=2}^{\infty} \lambda_i h_i(x) > x_0^* - x$. Now, $h_i(x) \geq h_{i+1}(x)$ iff $\left(\frac{1+\rho(1-2x)}{1-\rho(1-2x)} \right)^{i-2} \geq \frac{1-x_0^*}{x_0^*}$, and $\frac{1+\rho(1-2x)}{1-\rho(1-2x)} \geq \frac{1+\rho(1-2x_0^*)}{1-\rho(1-2x_0^*)}$ for all x in $(0, x_0^*)$. Thus, for any $i \geq i_{\max}$, we have $h_i(x) \geq h_{i+1}(x)$ for all $x \in (0, x_0^*)$.

$$\tilde{\lambda}_i = \begin{cases} \lambda_i, & i < i_{\max}, \\ \sum_{j=i_{\max}}^{\infty} \lambda_j, & i = i_{\max}, \\ 0, & i > i_{\max}. \end{cases}$$

Property $i)$ is immediate. To see that $ii)$ holds note that $\int\tilde{\lambda} - \int\lambda = \sum_{j=i_{\max}+1}^{\infty} \lambda_j (\frac{1}{i_{\max}+1} - \frac{1}{j}) > 0$. Finally, by construction $\sum_{i=2}^{\infty} \tilde{\lambda}_i h_i(x) \geq \sum_{i=2}^{\infty} \lambda_i h_i(x)$ for all $x \in (0, x_0^*)$ which proves property $iii)$. □

Note that one can always reduce the rate and increase the threshold of a degree sequence pair by altering ρ so that $\rho(x)$ increases for all $x \in (0, 1]$. Thus, (ii) can be made an equality if desired.

For any fixed rate $r \in (0, 1)$ it is usually easy to find a strictly positive lower bound for the optimal threshold of a given class of codes (by exhibiting an example), and an upper bound strictly less than 1 is always provided by the capacity formula $r \leq 1 - h(x_0^*)$, where $h(\cdot)$ is the binary entropy function. Theorem 1 then implies a strict upper bound on $i_{\max}$.

Consider now right-concentrated degree sequence pairs (λ, ρ). Note that a concentrated degree sequence ρ is uniquely determined by the parameter $\int\rho := \int_0^1 \rho(x)\, dx$. The following lemma states bounds for the parameter $\int\rho$ of the optimal (right-concentrated) degree sequence pair. The proof can be found in the Appendix.

Lemma 5. [Bounds on $\int\rho$] Let r be fixed and let $\underline{x_0^*}$ be a lower bound on the achievable threshold of a (right-concentrated) degree sequence pair. Let (λ, ρ) be the optimal (right-concentrated) degree sequence pair.

[Lower Bound] Let $\underline{\int\rho}$ be the unique positive root of the equation

$$x^2(1 - \underline{x_0^*}) + x\underline{x_0^*}(2 - r) - \underline{x_0^*}(1 - r) = 0.$$

Then $\int\rho \geq \underline{\int\rho}$.

[Upper Bound] Let $\overline{\lambda_2}$ be the smallest positive real root of the equation

$$(-1+r+2\underline{x_0^*}-2r\underline{x_0^*})\overline{\lambda_2}^2+(3+r(3-2\underline{x_0^*})-10\underline{x_0^*})\overline{\lambda_2}+2(-1+r+4\underline{x_0^*}+2r\underline{x_0^*})=0.$$

Then $\int\rho \leq \overline{\int\rho} := (1-r)\frac{\overline{\lambda_2}+2}{6}$.

We will now show that for a wide range of rates the optimum right-concentrated degree sequence pairs are also *left-concentrated*. We will establish this fact in the following manner. Assume that we are looking for right-concentrated degree sequence pairs which maximize X_0^* rather than the actual threshold x_0^*. For each such degree sequence pair it is an easy matter to check whether $x_0^* = X_0^*$. Luckily, it turns out that this is the case for all codes above a rate of roughly one-eighth. Recall that $X_0^* = \min\{\frac{1-\lambda_2\rho'(1)}{(\lambda'(1)-\lambda_2)\rho'(1)}, \tau\}$. The next two lemmas establish that each of the two respective quantities are maximized (for a given $\int\lambda$ and $\rho'(1)$) by left-concentrated degree sequence pairs.

Lemma 6. [Left Concentration Maximizes Stability Condition] For a fixed $\int\lambda$ and $\rho'(1)$ the quantity $\frac{1-\lambda_2\rho'(1)}{(\lambda'(1)-\lambda_2)\rho'(1)}$ is maximized by a concentrated degree sequence λ.

Proof. Assume $\lambda_i > 0$ and $\lambda_k > 0$ where $k-i>1$ and let $i<j<k$. Define $\eta_i = -(k-j)i$, $\eta_j = (k-i)j$, and $\eta_k = -(j-i)k$. It is easy to verify that $\eta_i+\eta_j+\eta_k = 0$ and $\frac{1}{i}\eta_i + \frac{1}{j}\eta_j + \frac{1}{k}\eta_k = 0$. Consider the polynomials given by $\lambda^\delta := \lambda + \delta\eta$ where $\eta_l = 0$ for $l \notin \{i,j,k\}$. For $\delta \in [0, \min\{-\eta_j/\lambda_j, -\eta_k/\lambda_k\}]$ we have $\lambda_i^\delta \geq 0$, $\sum_{i\geq 2}\lambda_i^\delta = 1$, and $\int\lambda^\delta = \int\lambda$. Now, $(i-1)\eta_i + (j-1)\eta_j + (k-1)\eta_k = -(k-i)(j-i)(k-j) < 0$, so $(\lambda^\delta)'(1)$ is decreasing in δ. This shows that $\lambda'(1)$ is minimized by (left) concentrated degree sequences. In the case $\lambda_2 = 0$ this proves the Lemma. Let us assume $\lambda_2 > 0$ and $i=2$. Then, letting λ depend on δ as above, we have

$$\frac{d}{d\delta}\frac{1-\lambda_2\rho'(1)}{(\lambda'(1)-\lambda_2)\rho'(1)} = \frac{-(1-\lambda_2\rho'(1))\frac{d}{d\delta}\lambda'(1) - (\lambda'(1)\rho'(1)-1)\frac{d}{d\delta}\lambda_2}{(\lambda'(1)-\lambda_2)^2\rho'(1)}.$$

Since $\frac{d}{d\delta}\lambda_2 < 0$ we see that $\frac{1-\lambda_2\rho'(1)}{(\lambda'(1)-\lambda_2)\rho'(1)}$ is increasing in δ. This shows that left concentrated degree sequences maximize this quantity. $\square$

Lemma 7. [Left Concentration Maximizes τ] For fixed $\int\lambda$, τ is maximized by a concentrated degree sequence λ.

Proof. Recall that τ is the smallest positive real root of the polynomial $p(x) := xp^+(x) + (x-1)p^-(x)$ and note that $p(x)$ is positive for x small enough. Now, the condition $xp^+(x)+(x-1)p^-(x) > 0$ is equivalent to $\frac{x}{1-x} > \frac{p^-(x)}{p^+(x)} = \frac{\sum_i \lambda_i r^{i-1}}{\sum_i \lambda_i(1-r)^{i-1}}$, where $r = \frac{1}{2}(1-\rho(1-2x))$. We will show that $\frac{\sum_i \lambda_i r^{i-1}}{\sum_i \lambda_i(1-r)^{i-1}}$ is simultaneously minimized over the whole range $r \in [0,1/2]$ by the unique concentrated degree sequence $\lambda(x)$ with parameter $\int\lambda$. This choice then clearly maximizes τ. First note that for $r=0$ and $r=1/2$ the value of $\frac{\sum_i \lambda_i r^{i-1}}{\sum_i \lambda_i(1-r)^{i-1}}$ is independent of λ. Thus, for $r \in (0,1/2)$ we want to find an optimal solution of the non-linear program

$$\min\{\frac{\sum_i \lambda_i r^{i-1}}{\sum_i \lambda_i(1-r)^{i-1}} | \lambda_i \geq 0; \sum_{i\geq 2}\lambda_i = 1; \sum_{i\geq 2}\lambda_i\frac{1}{i} = \int\lambda \}.$$

Consider the following linear program

$$\min\{\sum_{i\geq 2}\gamma_i r^{i-1}|\gamma_i \geq 0; \sum_{i\geq 2}\gamma_i(\textstyle\int\lambda - \frac{1}{i}) = 0; \sum_{i\geq 2}\gamma_i(1-r)^{i-1} = 1\}.$$

It is easy to check that if $\gamma(x)$ is an optimal solution of this linear program then $\lambda(x) := \gamma(x)/(\sum_{i\geq 2}\gamma_i)$ is an optimal solution to the above non-linear optimization problem. To solve the linear program let $i = \lfloor 1/\int\lambda \rfloor$. We claim that the optimal solution is given by

$$\begin{aligned}\gamma(x) \;:=\; & \frac{\int\lambda - \frac{1}{i+1}}{(\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i}x^{i-1} + \\ & \frac{\frac{1}{i} - \int\lambda}{(\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i}x^i.\end{aligned}$$

First note that the value of this feasible solution is equal to

$$\frac{(\int\lambda - \frac{1}{i+1})r^{i-1} + (\frac{1}{i} - \int\lambda)r^i}{(\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i}. \tag{8}$$

Now consider the dual program

$$\max\{y_2|y_1(\textstyle\int\lambda - \frac{1}{i}) + y_2(1-r)^{i-1} \leq r^{i-1}, i = 2, ...\}.$$

Define the tuple (y_1, y_2) by

$$\begin{aligned}y_1 &= \frac{r^i(1-r)^{i-1} - r^{i-1}(1-r)^i}{(\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i}, \\ y_2 &= \frac{(\int\lambda - \frac{1}{i+1})r^{i-1} + (\frac{1}{i} - \int\lambda)r^i}{(\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i}.\end{aligned}$$

Clearly, assuming that (y_1, y_2) is feasible, the value of this solution is equal to (8), the value of the solution of the primal problem. This will establish the optimality of the claimed solution. To prove the feasibility we need to verify for each $j \notin \{i, i+1\}$ that

$$\begin{aligned}&(r^i(1-r)^{i-1} - r^{i-1}(1-r)^i)(\textstyle\int\lambda - \frac{1}{j}) \\ &+((\textstyle\int\lambda - \frac{1}{i+1})r^{i-1} + (\frac{1}{i} - \int\lambda)r^i)(1-r)^{j-1} \\ &\leq ((\textstyle\int\lambda - \frac{1}{i+1})(1-r)^{i-1} + (\frac{1}{i} - \int\lambda)(1-r)^i)r^{j-1}.\end{aligned}$$

It is easy to check that this inequality holds for $r = 0$ and $r = 1/2$. We assume $r \in (0, 1/2)$ and $j \neq i, i+1$. After some manipulations (divide through by $r^{i-1}(1-r)^{i-1}$ and rearrange terms) the above inequality can be written as

$$\begin{aligned}0 \;\geq\; & \left(\tfrac{1}{i} - \textstyle\int\lambda\right)(r(1-r)^{j-i} - (1-r)r^{j-i}) \\ & + \left(\textstyle\int\lambda - \frac{1}{i+1}\right)((1-r)^{j-i} - r^{j-i}) \\ & - \left(\textstyle\int\lambda - \frac{1}{j}\right)(1-2r).\end{aligned} \tag{9}$$

The right hand side of (9) can be written as the sum of $\int\lambda\ ((1-r)^{j-i+1}-r^{j-i+1}-(1-2r))$ and a quantity independent of $\int\lambda$. If $j > i+1$, then $(1-r)^{j-i+1} - r^{j-i+1} - (1-2r)$ is convex for $r \in [0, 1/2]$ and equals 0 at both endpoints. Thus, in this case, the right hand side of (9) is maximized over $\int\lambda\ \in [\frac{1}{i+1}, \frac{1}{i}]$ by $\int\lambda\ = \frac{1}{i+1}$. In the case $j < i$ we have $(1-r)^{j-i+1} - r^{j-i+1} - (1-2r) < 0$ directly, hence, we can assume $\int\lambda\ = \frac{1}{i+1}$ in (9) in this case also.

Thus, setting $\int\lambda\ = \frac{1}{i+1}$ and multiplying by $i+1$, equation (9) simplifies to

$$0 \geq \tfrac{1}{i}(r(1-r)^{j-i} - (1-r)r^{j-i}) - \tfrac{j-i-1}{j}(1-2r)\,. \qquad (10)$$

Let us assume $j < i$. Noting that $-(r(1-r)^{j-i}-(1-r)r^{j-i}) = (1-2r)\sum_{l=0}^{i-j}\left(\frac{1}{r^{i-j-l}(1-r)^l}\right)$ we find that the inequality simplifies to

$$0 \geq -\frac{1}{i}\left(\sum_{l=0}^{i-j}\frac{1}{r^{i-j-l}(1-r)^l}\right) - \frac{j-i-1}{j}\,.$$

Now, $\sum_{l=0}^{i-j}\frac{1}{r^{i-j-l}(1-r)^l}$ is minimized over $r \in (0, 1/2]$ at $r = 1/2$ since $\frac{1}{r^l} + \frac{1}{(1-r)^l}$ and $\frac{1}{(r(1-r))^l}$ are both minimized at $r = 1/2$. At $r = 1/2$, (10) simplifies to

$$2^{i-j} \geq \frac{i}{j}\,,$$

which is clearly always fulfilled.

Let us assume $j > i+1$. Equation (10) can now be written as

$$0 \geq \tfrac{1}{i}r(1-r)(1-2r)\left(\sum_{l=0}^{j-i-2} r^{j-i-2-l}(1-r)^l\right) - \tfrac{j-i-1}{j}(1-2r)\,.$$

Now, $\sum_{l=0}^{j-i-2} r^{j-i-2-l}(1-r)^l \leq 1$ so the inequality will follow if we can show that $\frac{1}{i}r(1-r) - \frac{j-i-1}{j} \leq 0$. Since $r(1-r) \leq \frac{1}{4}$ it is sufficient to show that $\frac{1}{4i} - \frac{j-i-1}{j} \leq 0$. But $\frac{1}{4i} - \frac{j-i-1}{j} = \frac{1}{4ij}(j(1-4i)+4i(i+1)) \leq \frac{1}{4ij}((i+2)(1-4i)+4i(i+1)) = \frac{1}{4ij}(2-3i) < 0$. Thus, (10) is proved. □

We are now ready to find the optimal right-concentrated degree sequence pairs as follows. For a given rate r find a right-concentrated degree sequence pair of rate r with non-zero threshold. Now use this threshold in Lemma 5 to bound the range of $\int\rho$ of the optimal right-concentrated degree sequence pair. This bound implies a bound on the minimum and maximum right degree of the optimal right-concentrated degree sequence pair. Now, use Lemma 1 to bound the maximum degree of $\lambda(x)$. From the discussion above we know that the right-concentrated code which maximizes X_0^* is also left-concentrated. Given the bounds on the degrees there are now only a few possible choices for the at most two non-zero degrees of $\lambda(x)$ and $\rho(x)$. For each such choice there is a one-dimensional family of codes with the desired rate which we can parameterized by a real parameter a say. The optimal choice of a can now be determined analytically by observing that X_0^* reaches its maximum when $\frac{1-\lambda_2\rho'(1)}{(\lambda'(1)-\lambda_2)\rho'(1)} = \tau$. Finally, if for the code which maximizes X_0^* we have $x_0^* = X_0^*$ then a right-concentrated code with optimal threshold has been found. We note that empirically the condition $x_0^* = X_0^*$ is fulfilled for all optimal codes of rate at least one-eighth.

Example 9. [Rate one-half codes.] Applying the bounds above we find that the optimal rate one-half code is a member of the family of codes parameterized by

$$\begin{aligned}\lambda(x) &:= ax^2+(1-a)x^3,\\ \rho(x) &:= \frac{7a}{3}x^6+\frac{3-7a}{3}x^7,\end{aligned}$$

where $a \in [0, 3/7]$. For each code in this family we have $x_0^* = X_0^*$. Note that $\frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)} = \frac{3}{7(3-a)^2}$ is strictly increasing in the range $[0, 3/7]$ and takes on values in the range $[\frac{1}{21}, \frac{7}{108}]$. On the other hand one can check that the smallest real positive root τ is strictly decreasing in this range and $\tau(a=0) > \frac{1}{21}$ whereas $\tau(a=3/7) < \frac{1}{21}$. It follows that X_0^* takes on its maximal value for that unique a in the range $[0, 3/7]$ which satisfies

$$\tau = \frac{3}{7(3-a)^2}. \tag{11}$$

Clearly, this value can be determined to any desired degree of accuracy. We get $a = 0.1115$ and $x_0^* = 0.0513663$.

Whenever the condition $x_0^* = X_0^*$ fails, as it does for small rates, the optimal right-concentrated degree sequence pair has to be found by means of the following less efficient procedure. Rather than looking for the code which maximizes the threshold for a given rate we will maximize the rate for a given threshold. Given upper and lower bounds on $\int\rho$, an upper bound on the maximal degree of $\lambda(x)$, and a fixed threshold x_0^*, the right-concentrated degree sequence pair with maximal rate can be found by means of a one-dimensional family of infinite programs:

$$\sup_{\int\rho}\max\{\frac{1}{\int\rho}\sum_{i\geq 2}\lambda_i\frac{1}{i}|\lambda_i\geq 0;\sum_{i\geq 2}\lambda_i=1;\sum_{i\geq 2}\lambda_i h_i(x)\geq 0, x\in(0,x_0^*)\}.$$

By discretizing the ranges of $\int\rho$ and x the solution can be found to any desired degree of accuracy.

The result of this optimization is shown in Fig. 1 where the threshold is plotted as a function of the rate. The solid curve corresponds to the capacity formula $r = 1-h(x_0^*)$. The dashed curve corresponds to the optimal right-concentrated degree sequence pairs.

In the remainder of this section we will show that over a wide range of rates right-concentrated degree sequence pairs are indeed optimal.

Let $\rho(x)$ be a degree sequence, i.e., $\rho(x)$ is a function $\rho : [0,1] \to [0,1]$ of the form $\rho(x) := \sum_{i\geq 2}\rho_i x^{i-1}$ with non-negative coefficients and further $\rho(1) = 1$. Let $\mathcal{Q}$ be the set of all such degree sequences. Let $\mathcal{Q}_c$ be the set of all degree sequences with $\int\rho = c$. Denote the concentrated degree sequence with value $\int\rho = c$ by $\rho_c(x)$. As the next Lemma shows such degree sequence are extremal in a certain range.

Lemma 8. [Extremality of $\rho_c(x)$] $\rho_c(1-2x) \geq \rho(1-2x)$ for any $x \in [0, \frac{c}{2(1+c)}]$ and $\rho \in \mathcal{Q}_c$.

Proof: Look at the linear program

$$\min\{-\sum_{j\geq 2}\rho_j u^{j-1}|\rho_j\geq 0;\sum_{j\geq 2}\rho_j=1;\sum_{j\geq 2}\rho_j\frac{1}{j}=\int\rho\ \}$$

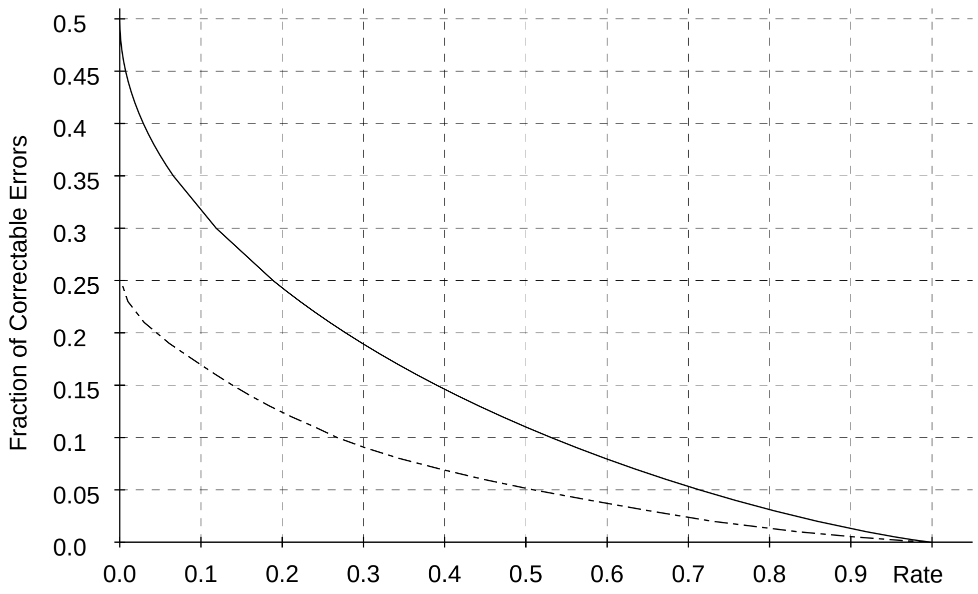

Figure 1: Optimal right-concentrated degree sequence pairs. The solid curve corresponds to the capacity formula $r = 1 - h(x_0^*)$ whereas the dashed curve corresponds to the optimal right-concentrated codes.

and its dual

$$\max\{y_1 + \textstyle\int\rho\ y_2 | y_1 + \frac{1}{j} y_2 \leq -u^{j-1}, j = 2, 3, \cdots\},$$

where $u \in [\frac{i}{i+1}, 1]$ and $i = \lfloor (\int\rho\ )^{-1} \rfloor$. We claim that the optimal solution of the primal problem is given by $\rho_j = 0$ for $j \notin \{i, i+1\}$ and

$$\rho_i = \frac{\int\rho\ - \frac{1}{i+1}}{\frac{1}{i} - \frac{1}{i+1}},$$
$$\rho_{i+1} = \frac{\frac{1}{i} - \int\rho}{\frac{1}{i} - \frac{1}{i+1}},$$

which has cost $\frac{(\frac{1}{i+1} - \int\rho\ )u^{i-1} + (\int\rho\ - \frac{1}{i})u^i}{\frac{1}{i} - \frac{1}{i+1}}$. To see this note that the tuple (y_1, y_2) given by

$$y_1 = \frac{\frac{1}{i+1}u^{i-1} - \frac{1}{i}u^i}{\frac{1}{i} - \frac{1}{i+1}},$$
$$y_2 = \frac{u^i - u^{i-1}}{\frac{1}{i} - \frac{1}{i+1}},$$

gives rise to the same cost. Hence, to prove optimality we need only show feasibility of this dual solution. Thus, for each $j \geq 2$ we require that the inequality $y_1 + \frac{1}{j} y_2 \leq -u^{j-1}$

holds. Now, after some manipulation, this inequality can be written in the form

$$i\left(1-\frac{i+1}{j}\right)-(i+1)\left(1-\frac{i}{j}\right)u+u^{j-i}\leq 0\,. \tag{12}$$

Note that for $j\in\{i,i+1\}$ the inequality (12) holds exactly, as expected. In general the expression on the left is a convex function of u, hence we need only verify (12) for $u=1$ and $u=\frac{i}{i+1}$. At $u=1$ the inequality (12) holds exactly, so we are left only to verify (12) for $u=\frac{i}{i+1}$. Substituting in this value (12) reduces to

$$i\geq j\left(\frac{i}{i+1}\right)^{j-i}. \tag{13}$$

Let us introduce the real parameter z in place of j, i.e., we consider $g(z):=z\left(\frac{i}{i+1}\right)^{z-i}$. The derivative of this quantity with respect to z is $\frac{d}{dz}g(z)=\left(\frac{i}{i+1}\right)^{z-i}(1-z\ln\frac{i+1}{i})$. Since $\frac{1}{i+1}<\ln\frac{1+i}{i}<\frac{1}{i}$, $g(z)$ is increasing on $[0,i]$ and decreasing on $[i+1,\infty)$. The claim now follows since as remarked above the inequality is fulfilled with equality for $j=i$ and $j=i+1$. □

The next lemma shows how this extremal property can be used to lower bound the threshold of some right-concentrated codes.

Lemma 9. [Right-Concentrated Codes] Let (λ,ρ) be a given degree sequence pair with $\rho\in\mathcal{Q}_c$ and with threshold $x_0^*(\lambda,\rho)$. Then $x_0^*(\lambda,\rho_c)\geq\min\{x_0^*(\lambda,\rho),\frac{c}{2(1+c)}\}$.

Proof. By assumption

$$x>x_0^*-x_0^*\lambda(\frac{1+\rho(1-2x)}{2})+(1-x_0^*)\lambda(\frac{1-\rho(1-2x)}{2})$$

for all $x\in(0,x_0^*)$. Let $\tilde{x_0^*}:=\min\{x_0^*(\lambda,\rho),\frac{c}{2(1+c)}\}$. By Lemma 8, $\rho_c(1-2x)\geq\rho(1-2x)$ for all $x\in[0,\tilde{x_0^*}]$. It follows that

$$\begin{aligned}
&\tilde{x_0^*}-\tilde{x_0^*}\lambda(\frac{1+\rho_c(1-2x)}{2})+(1-\tilde{x_0^*})\lambda(\frac{1-\rho_c(1-2x)}{2})\\
\leq\;&\tilde{x_0^*}-\tilde{x_0^*}\lambda(\frac{1+\rho(1-2x)}{2})+(1-\tilde{x_0^*})\lambda(\frac{1-\rho(1-2x)}{2})\\
\leq\;&x_0^*-x_0^*\lambda(\frac{1+\rho(1-2x)}{2})+(1-x_0^*)\lambda(\frac{1-\rho(1-2x)}{2})\\
<\;&x,
\end{aligned}$$

which shows that $x_0^*(\lambda,\rho_c)\geq\tilde{x_0^*}$. □

Lemma 10. [Optimality of Right-Concentrated Codes] Let (λ,ρ) be the optimal right-concentrated degree sequence pair of rate r and with threshold x_0^*. If

$$r<\frac{1-12x_0^*+16x_0^{*2}}{1-6x_0^*+8x_0^{*2}}$$

then (λ,ρ) is optimal.

Proof. Let $(\tilde{\lambda}, \tilde{\rho})$ be any optimal code of rate r and denote its threshold by $\tilde{x}_0^*$. If $x_0^* = \tilde{x}_0^*$ then we are done. So assume that $x_0^* < \tilde{x}_0^*$. Let $\underline{\int\tilde{\rho}}$ be any lower bound on the quantity $\int\tilde{\rho}$. Then using Lemma 9 we get

$$\begin{aligned} x_0^* &\geq x_0^*(\tilde{\lambda}, \tilde{\rho}_c) \\ &\geq \min\{\tilde{x}_0^*, \frac{c}{2(1+c)}\} \\ &\geq \min\{\tilde{x}_0^*, \frac{\underline{\int\tilde{\rho}}}{2(1+\underline{\int\tilde{\rho}}\,)}\}. \end{aligned}$$

Now, if $x_0^* < \frac{\underline{\int\tilde{\rho}}}{2(1+\underline{\int\tilde{\rho}}\,)}$ then this leads to a contradiction, showing that the optimal right-concentrated code is indeed globally optimal. But $x_0^* < \frac{\underline{\int\tilde{\rho}}}{2(1+\underline{\int\tilde{\rho}}\,)}$ is equivalent to $\underline{\int\tilde{\rho}} > \frac{2x_0^*}{1-2x_0^*}$. For $\underline{\int\tilde{\rho}}$ we can use the unique positive root of the equation $p(x) = 0$, where $p(x) := x^2(1-x_0^*) + x x_0^*(2-r) - x_0^*(1-r)$, as stated in Lemma 5. Since $p(0) < 0$ it follows that $x_0^* < \frac{\underline{\int\tilde{\rho}}}{2(1+\underline{\int\tilde{\rho}}\,)}$ iff $p(\frac{2x_0^*}{1-2x_0^*}) < 0$. Solving this equation for r results in the stated condition. □

Fig. 2 shows the curve (r, x_0^*) for the optimal right-concentrated code together with the optimality region marked in gray. We see that for rates roughly above 0.4, right-concentrated degree sequence pairs are optimal.

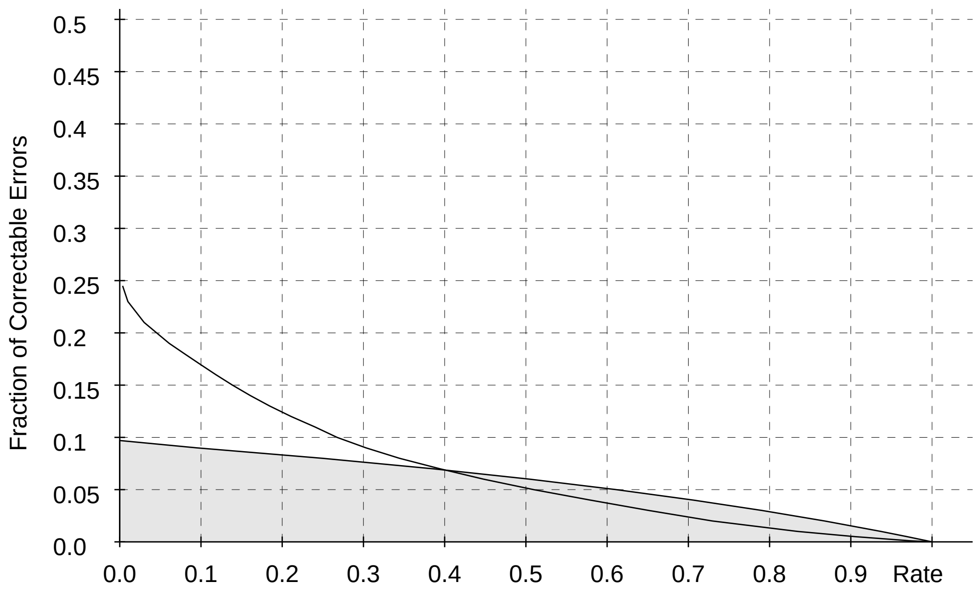

Figure 2: The curve (r, x_0^*) for the optimal right-concentrated code together with the optimality region marked in gray. For rates roughly above 0.4, right-concentrated degree sequence pairs are optimal.

4 Extension to Decoder with Erasures

In this section we outline how the previous analysis may be carried over to more complex decoders. We will focus on the decoder with erasures and the $(3,6)$-regular code.

Example 10. [Decoder with Erasures [4]] We assume that the received messages are elements of the alphabet $\{\pm 1\}$ and that they are the result of passing the transmitted word through a binary symmetric channel. Internally, the decoder uses messages in the extended alphabet $\{-1, 0, 1\}$, with 0 indicating an erasure. At a variable node the outgoing message along edge e is equal to sign $(wr + \sum_{\tilde{e} \neq e} m_{\tilde{e}})$, where r is the received value at the given variable node, w is a weight which is equal to 2 in the first iteration and 1 for all future iterations and the sum extends over all incoming messages at the given variable node except the one along edge e. In short, the outgoing message is the result of a simple *voting* scheme, whereby the received message has more weight in the first iteration reflecting its higher initial reliability. At a check node the outgoing message along edge e is the product of all incoming messages expect the incoming message along edge e. This is analogous to the case of Gallager's decoding algorithm A except that an outgoing message is now the erasure message if any of the involved incoming messages is an erasure message.

Specializing the equations in [4] to the case of a $(3,6)$-regular code we get the following set of equations: For $l = 1$ we have

$$\begin{aligned} p_1^- &= p_0^-(1 - (1 - q_0^- - q_0^0)^2), \\ p_1^0 &= p_0^-((1 - q_0^- - q_0^0)^2 - (q_0^-)^2) + (q_0^-)^2, \end{aligned}$$

whereas for $l > 1$ we have

$$\begin{aligned} p_l^- &= p_0^-((q_{l-1}^0)^2 + 2q_{l-1}^-(1 - q_{l-1}^-)) + (q_{l-1}^-)^2, && (14) \\ p_l^0 &= p_0^- 2q_{l-1}^0(1 - 2q_{l-1}^- - q_{l-1}^0) + 2q_{l-1}^- q_{l-1}^0, && (15) \end{aligned}$$

where

$$\begin{aligned} q_{l-1}^- &= \frac{1}{2}\left[(1 - p_{l-1}^0)^5 - (1 - 2p_{l-1}^- - p_{l-1}^0)^5\right], \\ q_{l-1}^0 &= 1 - (1 - p_{l-1}^0)^5. \end{aligned}$$

Without justification we apply the same method that succeeded in the case of the Gallager A decoder, i.e., we are trying to find fixed points of the equations in (14) and (15). Solving these equations for p_0^- and equating the result gives the equation

$$\frac{p^- - (q^-)^2}{(q^0)^2 + 2q^-(1 - q^-)} = \frac{p^0 - 2q^- q^0}{2q^0(1 - 2q^- - q^0)}.$$

This equation only involves the variables p^- and p^0. Further, in the range of interest there is a unique non-zero solution for p^0 for a given value of p^-. Given a fixed point pair (p^0, p^-) we can use lets say equation (14) to find the input error probability p_0^- for which the given pair is a fixed point. Fig. 3 shows p_0^- as a function of p^-. Note that

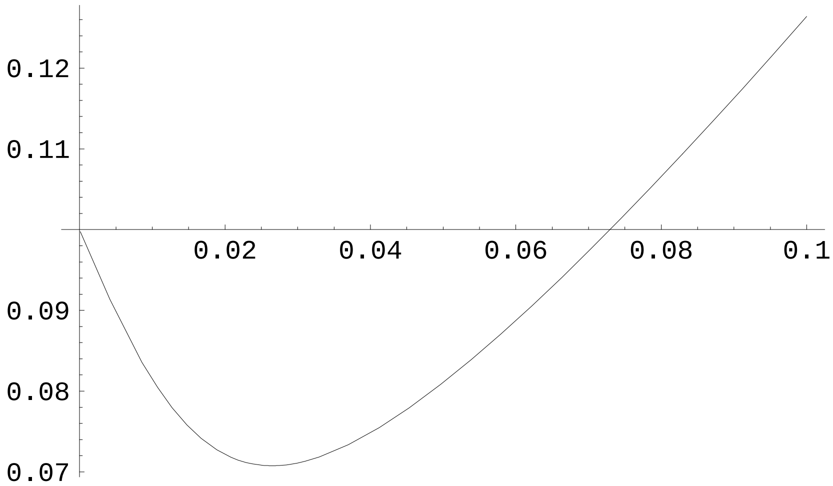

Figure 3: p_0^- as a function of p^- for the $(3,6)$-regular code. The smallest possible value of p_0^-, call it p_0^{-*}, is equal to 0.0708.

the smallest possible value of p_0^- is equal to $p_0^{-*} \sim 0.0708$. In words, if the initial error probability is less than p_0^{-*} then the equation system does not have fixed points (with non-zero values of p^0). If on the other hand the initial error probability is larger than p_0^{-*} then at least one such fixed point exists. It is now tempting to make the following

Conjecture 1. [Threshold of $(3,6)$-regular Code] The threshold of the $(3,6)$-regular code under the decoder with erasures is equal to $p_0^{-*} \sim 0.0708$.

Explicit numerical evaluations of the recursions support this conjecture. Unfortunately it seems no easy task to give a rigorous proof. For a discrete dynamical system in two dimensions the behavior of the system is not necessarily determined by the existence or lack of fixed points. The main problem lies in the fact that, contrary to the case of a belief propagation decoder, we do not know of a Ljapunov function for the given system. The analytic determination of the threshold (in terms of roots of some polynomial) is hence an open problem.

A Bounds and Extremality

In this Appendix we give the proof of Lemma 5.

Let (λ, ρ) be an optimal (right-concentrated) degree sequence pair with threshold x_0^*. Then

$$\frac{1-\lambda_2\rho'(1)}{\lambda'(1)\rho'(1)-\lambda_2\rho'(1)} \geq x_0^*. \tag{16}$$

Solving for $\rho'(1)$ we get

$$\rho'(1) \leq \frac{1}{\lambda_2(1-x_0^*)+x_0^*\lambda'(1)} \leq \frac{1}{x_0^*\lambda'(1)}. \tag{17}$$

The following bound will be used repeatedly.

Lemma 11. Let $\gamma(x)$ be a degree sequence. Then

$$\int\gamma \;\geq \frac{1}{\gamma'(1)+1}, \tag{18}$$

or equivalently

$$\gamma'(1) \geq \frac{1-\int\gamma}{\int\gamma}. \tag{19}$$

Proof.

$$\begin{aligned}
\int\gamma\;(\gamma'(1)+1) &= \left(\sum_{i\geq 2}\gamma_i i^{-1}\right)\left(\sum_{i\geq 2}\gamma_i(i-1)+1\right) \\
&= \left(\sum_{i\geq 2}\gamma_i i^{-1}\right)\left(\sum_{i\geq 2}\gamma_i i\right) \\
&\geq \left(\sum_{i\geq 2}\gamma_i\right)^2 \\
&= 1\,,
\end{aligned}$$

where for the third line we used the Cauchy-Schwartz inequality. □

We are now ready for the

Proof of Lemma 5. We start with the lower bound. Let (λ,ρ) be the optimal (right-concentrated) degree sequence pair. Making repeated use of Lemma 11 and using equation (17) we get

$$\begin{aligned}
\frac{1-\int\rho}{\int\rho} &\leq \rho'(1) \\
&\leq \frac{1}{x_0^*\lambda'(1)} \\
&\leq \frac{1}{x_0^*\frac{1-\int\lambda}{\int\lambda}} \\
&= \frac{1}{x_0^*\frac{1-\frac{1}{1-r}\int\rho}{\frac{1}{1-r}\int\rho}} \\
&= \frac{\int\rho}{x_0^*(1-r-\int\rho\;)}.
\end{aligned}$$

Letting $x := \int\rho\;$ we get the said equation.

For the upper bound we start by solving equation (16) for $\overline{\lambda_2}$. We get

$$\lambda_2 \leq \frac{1-x_0^*\lambda'(1)\rho'(1)}{\rho'(1)(1-x_0^*)}.$$

It follows that

$$\overline{\lambda_2} = \frac{1-x_0^*\underline{\lambda'(1)\rho'(1)}}{\underline{\rho'(1)}(1-x_0^*)}.$$

Given an upper bound on λ_2 we can upper bound $\int\lambda$ and lower bound $\lambda'(1)$ by

$$\overline{\int\lambda} = \frac{\overline{\lambda_2}+2}{6}$$

and

$$\underline{\lambda'(1)} = 2 - \overline{\lambda_2}.$$

Next the rate condition implies that

$$\overline{\int\rho} = (1-r)\overline{\int\lambda}\,.$$

Finally,

$$\underline{\rho'(1)} = \frac{1-\overline{\int\rho}}{\overline{\int\rho}}.$$

Combining these equations results in the said formula. □

B Acknowledgments

We would like to thank Amin Shokrollahi for interesting and useful discussions.

References

[1] R. G. Gallager, *Low-Density Parity-Check Codes.* Cambridge, Massachusetts: M.I.T. Press, 1963.

[2] D. J. C. MacKay, "Good error correcting codes based on very sparse matrices," *IEEE Trans. on Information Theory*, vol. 45, pp. 399–431, 1999.

[3] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman, and V. Stemann, "Practical loss-resilient codes," in *Proceedings of the* 29*th annual ACM Symposium on Theory of Computing*, pp. 150–159, 1997.

[4] T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message-passing decoding." submitted IEEE IT.

[5] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of provably good low-density parity check codes." submitted IEEE IT.

[6] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann Publishers, 1986.

[7] R. Descartes, Geometrie *1636, in A Source Book in Mathematics.* Cambridge: Harvard Univ. Press, 1969.

[8] P. Henrici, *Applied and Computational Complex Analysis, Vol. 1.* New York: John Wiley & Sons, 1974.

[9] A. Schrijver, *Theory of linear and integer programming.* Wiley, 1986.