Spatially-Coupled Precoded Rateless Codes

Kosuke Sakata, Kenta Kasai and Kohichi Sakaniwa
Dept. of Communications and Integrated Systems, Tokyo Institute of Technology, 152-8550 Tokyo, Japan.
Email: {sakata,kenta,sakaniwa}@comm.ss.titech.ac.jp

***Abstract*— Raptor codes are rateless codes that achieve the capacity on the binary erasure channels. However the maximum degree of optimal output degree distribution is unbounded. This leads to a computational complexity problem both at encoders and decoders. Aref and Urbanke investigated the potential advantage of universal achieving-capacity property of proposed spatially-coupled (SC) low-density generator matrix (LDGM) codes. However the decoding error probability of SC-LDGM codes is bounded away from 0. In this paper, we investigate SC-LDGM codes concatenated with SC low-density parity-check codes. The proposed codes can be regarded as SC Hsu-Anastasopoulos rateless codes. We derive a lower bound of the asymptotic overhead from stability analysis for successful decoding by density evolution. The numerical calculation reveals that the lower bound is tight. We observe that with a sufficiently large number of information bits, the asymptotic overhead and the decoding error rate approach 0 with bounded maximum degree.**

I. Introduction

Spatially-coupled (SC) low-density parity-check (LDPC) codes attract much attention due to their capacity-achieving performance under low-latency memory-efficient sliding-window belief propagation (BP) decoding. The studies on SC-LDPC codes date back to the invention of convolutional LDPC codes by Felström and Zigangirov [1]. Lentmaier *et al.*! observed that the BP threshold of regular SC-LDPC codes coincides with the maximum a posterior (MAP) threshold of the underlying block LDPC codes with a lot of accuracy by density evolution [2]. Kudekar *et al.* proved that SC-LDPC codes achieve the MAP threshold of BEC [3] and the binary-input memoryless output-symmetric (BMS) channels [4] under BP decoding.

Rateless codes are a class of erasure-recovering codes which produce limitless sequence of encoded bits from k information bits so that receivers can recover the k information bits from arbitrary $(1+\alpha)k/(1-\epsilon)$ received symbols from BEC(ϵ). We denote *overhead* by α. Designing rateless codes with vanishing overhead is desirable, which implies the codes achieve the capacity of BEC(ϵ). LT codes [5] and raptor codes [6] are rateless codes that achieve vanishing overhead $\alpha \to 0$ in the limit of large information size over the BEC. By a nice analogy between the BEC and the packet erasure channel (e.g., Internet), rateless codes have been successfully adopted by several industry standards.

A raptor code can be viewed as concatenation of an outer high-rate LDPC code and infinitely many single parity-check codes of length d, where d is chosen randomly with probability Ω_d for $d \geq 1$. Raptor codes need to have unbounded maximum degree d for $\Omega_d \neq 0$. This leads to a computation complexity problem both at encoders and decoders.

The authors presented empirical results in [7] showing that SC MacKay-Neal (MN) codes and SC Hsu-Anastasopoulos (HA) codes achieve the capacity of BEC with bounded maximum degree. Recently a proof for SC-MN codes are given in [8]. It was observed that the SC-MN codes and SC-HA codes have the BP threshold close to the Shannon limit in [9] over BMS channels.

Aref and Urbanke [10] investigated the potential advantage of universal achieving-capacity property of SC low-density generator matrix (LDGM) codes. They observed that the decoding error probability steeply decreases with overhead $\alpha = 0$ with bounded maximum degree over various BMS channels. However the decoding error probability was proved to be bounded away from 0 with bounded maximum degree for any α. This is explained from the fact that there are a constant fraction of bit nodes of degree 0.

In this paper, we investigate SC-LDGM codes concatenated with SC-LDPC codes. The proposed codes can be regarded as SC-HA rateless codes. We derive a lower bound of the asymptotic overhead from stability analysis for successful decoding by density evolution. The numerical calculation reveals that the lower bound is tight. We observe that with a sufficiently large number of information bits, the asymptotic overhead and the decoding error rate approach 0 with bounded maximum degree.

II. Encoder and Decoder

A. Encoder

Let k denote the number of information bits. We define a (d_l, d_r, d_g, L, w) code for $d_l \geq 2, d_r \geq 2, d_g \geq 2$ as follows. The (d_l, d_r, d_g, L, w) code are defined on L sections from 0 to $L-1$. Each section has M pre-coded bits. Note that, in [3], $2L+1$ sections $[-L, +L]$ were considered. Instead, for the sake of simplicity, we consider L sections in $[0, L-1]$. First, the k information bits are pre-coded with (d_l, d_r, L, w) codes [3] into LM bits $x(0,0), \dots, x(L-1, M-1)$. In this paper, we assume that the bits in the i-th section for $i \in [0, L-1]$ are transmitted and the bits in other sections are shortened. Namely, the shortened bits are set to 0 and are not transmitted. Let $R_{\mathrm{pre}}(L)$ denote the design coding rate of (d_l, d_r, w, L) codes. In [3], $R_{\mathrm{pre}}(L)$ is given by

$$R_{\mathrm{pre}}(L) = 1 - \frac{d_l}{d_r} - \frac{d_l}{d_r}\frac{w - 1 - 2\sum_{i=1}^{w-1}(i/w)^{d_r}}{L}$$

$$\overset{L\to\infty}{=} 1 - \frac{d_l}{d_r}.$$

It follows that $k = R_{\mathrm{pre}}(L)LM$.

After encoding the k bits into LM coded bits by pre-code, the LM pre-coded bits further will be encoded by an inner code as follows. Repeat the following procedure endlessly for $t \in [1, \infty)$.

1) Choose a section $i^{(t)} \in [0, L+w-2]$ uniformly at random from $L+w-1$ sections.
2) Choose d_g section shifts $j_1^{(t)}, \dots, j_{d_g}^{(t)} \in [0, w-1]$ with repetition uniformly at random.
3) Choose d_g bit-indices $l_1^{(t)}, \dots, l_{d_g}^{(t)} \in [0, M-1]$ with repetition uniformly at random.
4) Add d_g bits and transmit the sum as

$$x(i^{(t)} - j_1^{(t)}, l_1^{(t)}) + \cdots + x(i^{(t)} - j_{d_g}^{(t)}, l_{d_g}^{(t)}). \quad (1)$$

B. Decoder

Assume that transmission takes place over BEC(ϵ) and we have n received symbols $y^{(1)}, \dots, y^{(n)}$ each of which is 0, 1 or '?'. Define the *overhead* α as

$$\alpha = \frac{n}{k}(1-\epsilon) - 1.$$

In this setting, we have $(1+\alpha)k = n(1-\epsilon)$ unerased received symbols. Independence of the coding scheme ensures that we can assume, without loss of generality, that time indices of n received symbols are arbitrary. For simplicity, we assume that the receiver receives n symbols at time $t = 1, \dots, n$ without loss of generality.

We assume that the decoder knows $i^{(t)}$, d_g section shifts $j_1^{(t)}, \dots, j_{d_g}^{(t)}$ and bit-indices $l_1^{(t)}, \dots, l_{d_g}^{(t)}$ in (1) for each received symbol at time $t = 1, \dots, n$. From these information and the knowledge of the precode, one can construct a factor graph for sum-product decoding [11]. The factor graph consists of LM variable nodes (bit nodes) $x(0,0), \dots, x(L-1, M-1)$ and $(1 - R_{\mathrm{pre}}(L))LM$ parity-check factor nodes (check nodes) of pre-code and factor nodes (channel nodes) of factor

$$\mathbf{1}[x(i^{(t)} - j_1^{(t)}, j_1^{(t)}) + \cdots + x(i^{(t)} - j_{d_g}^{(t)}, j_{d_g}^{(t)}) = y^{(t)}] \quad (2)$$

for $t = 1, \dots, n$, where $\mathbf{1}[\cdot]$ is defined as 1 if the argument is true and 0 otherwise. We say that the factor node of factor (2) is in the section $i^{(t)}$.

III. Performance Analysis

In this section, we investigate the performance of the coupled rateless codes and derive a bound.

A. Performance Analysis by Density Evolution

In this subseciton, we derive the density evolution update equation. The following lemma clarifies the degree distributions of inner codes.

Lemma 1: Let Λ_d be the probability that a bit node has d neighboring channel nodes. Let β be the average number of channel nodes adjacent to a bit node. In the limit of large M, we have

$$\beta = \frac{d_g}{1-\epsilon} \frac{LR_{\mathrm{pre}}(L)(1+\alpha)}{L+w-1}, \quad (3)$$

$$\sum_{d\geq 0} \Lambda_d x^d = e^{-\beta(1-x)} = \sum_{d\geq 0} \frac{\beta^d e^{-\beta}}{d!} x^d.$$

Proof: Let N denote the average number of channel nodes per section. There are $L+w-1$ sections containing channel nodes. We have n channel nodes in total.

$$N = \frac{n}{L+w-1} = \frac{1}{1-\epsilon} \frac{(1+\alpha)k}{L+w-1} = \frac{1}{1-\epsilon} \frac{(1+\alpha)R_{\mathrm{pre}}(L)LM}{L+w-1},$$

where we used $k = R_{\mathrm{pre}}(L)LM$. Recalling that β is the average number of channel nodes adjacent to a bit node, we have

$$\beta = \frac{d_g N}{M}.$$

Equation (3) immediately follows from this. Each section has N channel nodes of degree d_g, in other words, we have $d_g N$ edges in each section. Let Λ_d denote the probability that a bit node in the i-th section has d channel nodes within sections from i to $i+w-1$. Since each channel node is generated independently, the probability Λ_d follows a binomial distribution as follows.

$$\Lambda_d = \binom{d_g N}{d} \left(\frac{1}{M}\right)^d \left(1 - \frac{1}{M}\right)^{d_g N - d}$$

The probability generating function of Λ_d is given as follows.

$$\Lambda(x) := \sum_{d\geq 0} \Lambda_d x^d = \left(\frac{x}{M} + 1 - \frac{1}{M}\right)^{d_g N} \overset{M\to\infty}{=} \exp[-\beta(1-x)] = \sum_{d\geq 0} \frac{\beta^d e^{-\beta}}{d!} x^d.$$

This implies $\Lambda_d = \frac{\beta^d e^{-\beta}}{d!}$ in the limit of $M \to \infty$. In other words, the degree d follows the Poisson distribution of average β. □

Let us describe density evolution update equations. Let $p_i^{(\ell)}$ and $s_i^{(\ell)}$ be the erasure probability of messages sent from bit nodes in the i-th section to check nodes and channel nodes, respectively, at the ℓ-th iteration of BP decoding of (d_l, d_r, d_g, L, w) codes in the limit of large M. The density evolution [12] gives update equations for $p_i^{(\ell)}$ and $s_i^{(\ell)}$ as follows. For $i \notin [0, L-1]$, $p_i^{(\ell)} = s_i^{(\ell)} = 0$. For $i \in [0, L-1]$, $p_i^{(0)} = s_i^{(0)} = 1$, and for $\ell \geq 0$,

$$p_i^{(\ell+1)} = \Big(\frac{1}{w}\sum_{j=0}^{w-1}\big(1 - (1 - \frac{1}{w}\sum_{k=0}^{w-1} p_{i+j-k}^{(\ell)})^{d_r - 1}\big)\Big)^{d_l - 1} \cdot \Lambda\Big(\frac{1}{w}\sum_{j=0}^{w-1}\big(1 - (1-\epsilon)(1 - \frac{1}{w}\sum_{k=0}^{w-1} s_{i+j-k}^{(\ell)})^{d_g - 1}\big)\Big),$$

$$s_i^{(\ell+1)} = \Big(\frac{1}{w}\sum_{j=0}^{w-1}\big(1-(1-\frac{1}{w}\sum_{k=0}^{w-1}p_{i+j-k}^{(\ell)})^{d_r-1}\big)\Big)^{d_l} \cdot \lambda\Big(\frac{1}{w}\sum_{j=0}^{w-1}\big(1-(1-\epsilon)(1-\frac{1}{w}\sum_{k=0}^{w-1}s_{i+j-k}^{(\ell)})^{d_g-1}\big)\Big),$$

where $\lambda(x) = \frac{\Lambda'(x)}{\Lambda'(1)} = \exp[-\beta(1-x)] = \Lambda(x)$.

Let $\mathbb{P}_{\mathrm{b}}^{(\ell)}$ be the decoding error probability at the ℓ-th iteration of BP decoding given as follows.

$$\mathbb{P}_{\mathrm{b}}^{(\ell)} := \frac{1}{L}\sum_{i=1}^{L} p_i^{(\ell)}.$$

Definition 1: One can easily check $\mathbb{P}_{\mathrm{b}}^{(\ell)}$ has its limit $\mathbb{P}_{\mathrm{b}}^{(\infty)}(L) := \lim_{\ell\to\infty}\mathbb{P}_{\mathrm{b}}^{(\ell)}(L)$ since $\mathbb{P}_{\mathrm{b}}^{(\ell)}$ is decreasing in ℓ. We define *overhead threshold* α_L^* and its corresponding β_L^* as follows.

$$\alpha_L^* := \inf\{\alpha > 0 \mid \mathbb{P}_{\mathrm{b}}^{(\infty)}(L) = 0\},$$
$$\beta_L^* := \inf\{\beta > 0 \mid \mathbb{P}_{\mathrm{b}}^{(\infty)}(L) = 0\}.$$

We say (d_l, d_r, d_g, L, w) codes *achieve the capacity of* BEC(ϵ) if

$$\limsup_{L\to\infty} \alpha_L^* = 0.$$

Discussion 1: We will explain why we exclude the case $d_g = 1$. Assume $d_g = 1$. The density evolution update equations can be reduced as follows.

$$p_i^{(\ell+1)} = \begin{cases} \Lambda(\epsilon)\Big(\frac{1}{w}\sum_{j=0}^{w-1}\big(1-(1-\frac{1}{w}\sum_{k=0}^{w-1}p_{i+j-k}^{(\ell)})^{d_r-1}\big)\Big)^{d_l-1} & (i\in[0,L-1]), \\ 0 & (i\notin[0,L-1]). \end{cases}$$

This is equivalent to the density evolution update equation of the precode that is a (d_l, d_r, w, L) code transmitted over BEC($\Lambda(\epsilon)$) [3]. If the error probability goes to 0, $\Lambda(\epsilon)$ has to be less than the Shannon limit $\Lambda(\epsilon) = e^{-\beta_L^*(1-\epsilon)} < 1 - R_{\mathrm{pre}}(L)$. It follows that β_L^* is bounded as follows.

$$\beta_L^* > \frac{1}{1-\epsilon}\ln\frac{1}{1-R_{\mathrm{pre}}(L)}.$$

From (4) we have

$$\alpha_L^* > \frac{L+w-1}{LR_{\mathrm{pre}}(L)}\ln\frac{1}{1-R_{\mathrm{pre}}(L)} - 1 \overset{L\to\infty}{=} \frac{d_r}{d_r-d_l}\ln\frac{d_r}{d_l} - 1 > 0.$$

This implies the $(d_l, d_r, d_g = 1, L, w)$ codes do not achieve the capacity of BEC(ϵ). This is the reason why we exclude the case $d_g = 1$ in this paper.

Lemma 2: The (d_l, d_r, d_g, L, w) codes achieve the capacity of BEC(ϵ) if and only if

$$\limsup_{L\to\infty}\beta_L^* = \frac{d_g}{1-\epsilon}\Big(1-\frac{d_l}{d_r}\Big).$$

Proof: This is straightforward from (3), we have

$$\beta_L^* = \frac{d_g}{1-\epsilon}R_{\mathrm{pre}}(L)\frac{L}{L+w-1}(1+\alpha_L^*) = \frac{d_g}{1-\epsilon}\Big(1-\frac{d_l}{d_r}\Big) \quad (L\to\infty). \tag{4}$$

□

B. Performance Bound by Stability Analysis

In the following theorem, we derive a lower bound of overhead threshold α_L^*.

Theorem 1: For $(d_l = 2, d_r, d_g, L, w)$ codes, if $\mathbb{P}_{\mathrm{b}}^{(\infty)}(L) = 0$ then there exist $\underline{\alpha}_L^*$ and $\underline{\beta}_L^*$ such that

$$\alpha_L^* \geq \underline{\alpha}_L^*, \quad \beta_L^* \geq \underline{\beta}_L^*,$$
$$\lim_{L\to\infty}\underline{\alpha}_L^* = \max\Big[\frac{\ln(d_r-1)}{d_g(1-2/d_r)} - 1, 0\Big],$$
$$\lim_{L\to\infty}\underline{\beta}_L^* = \max\Big[\frac{\ln(d_r-1)}{1-\epsilon}, \frac{d_g}{1-\epsilon}\Big(1-\frac{d_l}{d_r}\Big)\Big].$$

Proof: Let P_L denote an $L\times L$ matrix whose (i,j) entry is $\frac{\partial p_i^{(\ell+1)}}{\partial p_j^{(\ell)}}$. As we will see, this does not depend on ℓ. Let $\rho(P_L)$ denote the spectral radius of P_L. We will derive a lower bound of $\rho(P_L)$.

Some calculation reveals that at $\boldsymbol{p}^{(\ell)} = \boldsymbol{s}^{(\ell)} = \mathbf{0}$ for $d_l = 2$.

$$\frac{\partial p_i^{(\ell+1)}}{\partial p_j^{(\ell)}} = \frac{(d_r-1)\lambda(\epsilon)}{w^2}\frac{\partial}{\partial p_j^{(\ell)}}\sum_{l=0}^{w-1}\sum_{k=0}^{w-1}p_{i+l-k}^{(\ell)} = \begin{cases} \frac{w-|i-j|}{w^2}(d_r-1)\lambda(\epsilon) & (|i-j|\leq w) \\ 0 & (|i-j|>w) \end{cases} \tag{5}$$

and $\frac{\partial p_i^{(\ell+1)}}{\partial p_j^{(\ell)}} = 0$ for $d_l > 2$. It holds that for $d_l \geq 2$,

$$\frac{\partial p_i^{(\ell+1)}}{\partial s_j^{(\ell)}} = \frac{\partial s_i^{(\ell+1)}}{\partial p_j^{(\ell)}} = \frac{\partial s_i^{(\ell+1)}}{\partial s_j^{(\ell)}} = 0.$$

at $\boldsymbol{p}^{(\ell)} = \boldsymbol{s}^{(\ell)} = \mathbf{0}$. We drop ℓ since (5) is independent of ℓ.

From (5), we can see that P_L is a *positive band matrix of width* w, which is defined in Definition 4 in Appendix. Since P_L is a positive band matrix of width w, one can see that P_L is irreducible from Lemma 4 in Appendix. Let $\lambda_1, \ldots, \lambda_L$ be the eigenvalues of P_L, recall that $\rho(P_L)$ is the spectral radius of P_L. We have

$$\rho(P_L) := \max_i(|\lambda_i|).$$

Since P_L is symmetric, the eigenvalues are real.

Let $\lambda_1 > \ldots > \lambda_L$ be the eigenvalues of P_L. Perron-Frobenius theorem [13] asserts that the eigenvalue that gives the spectral radius of a non-negative irreducible matrix is positive. Since P_L is non-negative symmetric irreducible matrix, the eigenvalue that gives spectral radius of P_L is positive. Then we have

$$\rho(P_L) = \lambda_1. \tag{6}$$

For $\delta > 0$, we define $\beta := \beta_L^* + \delta$. Since $\beta > \beta_L^*$, it follows $\mathbb{P}_{\mathrm{b}}^{(\infty)}(L) = 0$. From (6), we have for $\forall \boldsymbol{x} \in \mathbb{R}^L \setminus \{\mathbf{0}\}$,

$$\begin{aligned}1 > \rho(P_L) &\overset{(a)}{=} \max_{\boldsymbol{x}\in\mathbb{R}^L:\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\mathsf{T} P_L \boldsymbol{x}}{\boldsymbol{x}^\mathsf{T}\boldsymbol{x}} \geq \frac{\mathbf{1}^\mathsf{T} P_L \mathbf{1}}{\mathbf{1}^\mathsf{T}\mathbf{1}} \\ &= (d_r-1)e^{-\beta(1-\epsilon)} \frac{w^2 L - (w-1)w(w+1)/3}{w^2 L} \\ &\overset{L\to\infty}{=} (d_r-1)e^{-\beta_L^*(1-\epsilon)}, \end{aligned} \tag{7}$$

where we used [14, Theorem 4.2.2] for (a). Solving β from this inequality, we obtain

$$\beta > \frac{1}{1-\epsilon} \ln\Big[(d_r-1)\big(1 - \frac{(w-1)(w+1)}{3wL}\big)\Big]. \tag{8}$$

$\lim_{\delta\to 0} \beta = \beta_L^*$ denote that $\beta_L^* \geq$ RHS of (8). A trivial lower bound $\alpha_L^* \geq 0$ is true, since we can not surpass the capacity. From this and (4), it follows that

$$\beta_L^* \geq \max\Big[\text{RHS of (8)}, \frac{d_g}{1-\epsilon} R_{\mathrm{pre}}(L) \frac{L}{L+w-1}\Big] =: \underline{\beta}_L^*$$
$$\alpha_L^* \geq \frac{\underline{\beta}_L^*(1-\epsilon)(L+w-1)}{d_g L R_{\mathrm{pre}}(L)} - 1 =: \underline{\alpha}_L^*$$

In the limit of large L, we have

$$\lim_{L\to\infty} \underline{\beta}_L^* = \max\Big[\frac{\ln(d_r-1)}{1-\epsilon}, \frac{d_g}{1-\epsilon}\Big(1 - \frac{d_l}{d_r}\Big)\Big],$$
$$\lim_{L\to\infty} \underline{\alpha}_L^* = \max\Big[\frac{d_r \ln(d_r-1)}{d_g(d_r-2)} - 1, 0\Big].$$

This concludes Theorem 1.

Discussion 2: For $L \geq 2w-1$, P_L have entries taking value from 1 to w. From [14, Lemma 5.6.10], we can bound $\rho(P_L)$ as follows.

$$\begin{aligned}\rho(P_L) \leq \|P_L\|_1 &:= \max_{1\leq i\leq L} \sum_{j=1}^{L} |(P_L)_{i,j}| \\ &= (d_r-1)e^{-\beta(1-\epsilon)} \frac{1}{w^2}\Big(w + 2\sum_{i=1}^{w-1} i\Big) \\ &= (d_r-1)e^{-\beta(1-\epsilon)}\end{aligned}$$

From this, we can see that the bound (7) is tight for large L. □

Corollary 1: For capacity-achieving $(d_l = 2, d_r, d_g, L, w)$ codes have to satisfy

$$d_g \geq \frac{d_r \ln(d_r-1)}{d_r-2}. \tag{9}$$

This condition is not satisfied for $d_r = 2$ or $d_g = 2$.

Proof: From Definition 1, capacity-achieving codes satisfy $\underline{\alpha}_L^*$ goes to 0 in the limit of large L. To be precise,

$$\lim_{L\to\infty} \underline{\alpha}_L^* = \max\Big[\frac{d_r \ln(d_r-1)}{d_g(d_r-2)} - 1, 0\Big] = 0.$$

The inequality (9) immediately follows from this. □

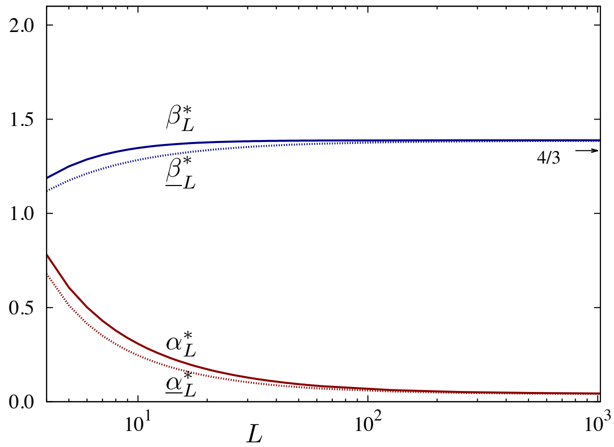

Fig. 1. The asymptotic overhead α_L^* and the average degree of β_L^* and their lower bounds $\underline{\alpha}_L^*$ and $\underline{\beta}_L^*$ of $(d_l = 2, d_r = 3, d_g = 2, L, w = 2)$ codes over BEC(ϵ=0.5). The asymptotic overhead threshold α_L^* does not converge to 0 since the codes do not satisfy the condition of Corollary 1. Figure suggests the lower bounds are tight for large L.

IV. Decoding Performance

In this section, we demonstrate the decoding performance of the (d_l, d_r, d_g, L, w) codes.

Figure 1 shows convergence the overhead threshold α_L^* and β_L^* and their lower bounds $\underline{\alpha}_L^*$ and $\underline{\beta}_L^*$ of $(d_l = 2, d_r = 3, d_g = 2, L, w = 2)$ codes over BEC(ϵ=0.5). The codes do not satisfy the condition of Corollary 1. This explains why α_L^* does not converge to 0 and β_L^* does not converge to 4/3 which is given in Lemma 2 as the limiting value of capacity-achieving codes. We observe that α_L^* approaches $\underline{\alpha}_\infty^* = \frac{3\ln(2)-2}{2} \simeq 0.03972$ and β_L^* approaches $\underline{\beta}_\infty^* = 2\ln(2) \simeq 1.38629$ which suggest the lower bounds are tight for large L.

Figure 2 shows convergence the asymptotic overhead threshold α_L^* and the average degree of β_L^* and their lower bounds $\underline{\alpha}_L^*$ and $\underline{\beta}_L^*$ of $(d_l = 2, d_r = 3, d_g = 3, L, w = 2)$ codes over BEC(ϵ=0.5). The codes satisfy the condition of Corollary 1. Thought this does not necessarily ensure α_L^* approaches 0 and β_L^* approaches 2 which is given in Lemma 2 as the limiting value of capacity-achieving codes, this is likely the case. We observe that α_L^* approaches $\underline{\alpha}_\infty^* = 0$ and β_L^* approaches $\underline{\beta}_\infty^* = 2$, which suggest the lower bounds are tight for large L.

Figure 3 compares approaching speed of overhead threshold α_L^* of $(d_l = 2, d_r, d_g = 3, L, w = 2)$ codes with $d_r \in \{3, 4, 14, 15, 20, 30\}$ over BEC(ϵ=0.5). The codes of $d_r \leq 14$ satisfy the condition of Corollary 1, while the codes of $d_r > 14$ do not. The fastest approaching speed is attained at $d_r = 14$.

V. Conclusion

We propose spatially-coupled precoded regular rateless codes. We have derived a lower bound $\underline{\alpha}_L^*$ of asymptotic overheads threshold α_L^*. The numerical calculation of density evolution shows that the bound is tight for large coupling number L and asymptotic overheads threshold α_L^* goes to 0 for large L with bounded density. The possible future work

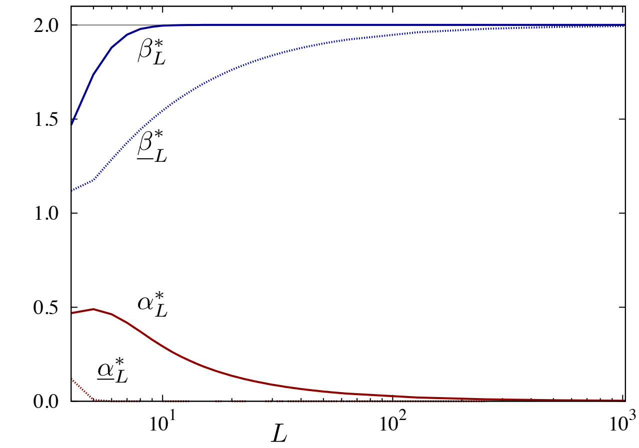

Fig. 2. The asymptotic overhead α_L^* and the average degree of β_L^* and their lower bounds $\underline{\alpha}_L^*$ and $\underline{\beta}_L^*$ of $(d_l = 2, d_r = 3, d_g = 3, L, w = 2)$ codes over BEC(ϵ=0.5). The codes satisfy the condition of Corollary 1. We observe that α_L^* approaches $\underline{\alpha}_\infty^* = 0$ and β_L^* approaches $\underline{\beta}_\infty^* = 2$, which suggest the lower bounds are tight for large L.

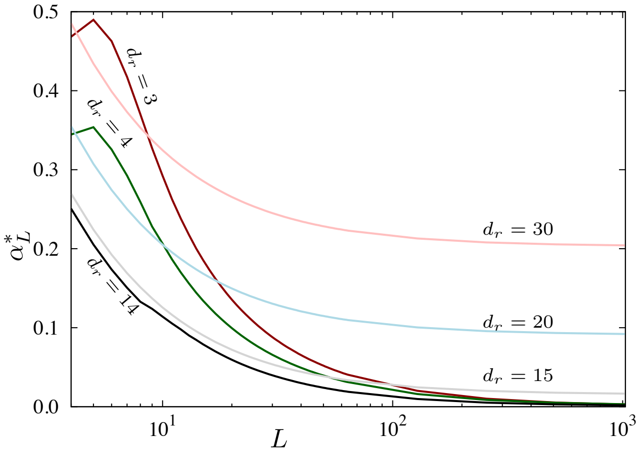

Fig. 3. Comparison of approaching speed of overhead threshold α_L^* of $(d_l = 2, d_r, d_g = 3, L, w = 2)$ codes with $d_r \in \{3, 4, 14, 15, 20, 30\}$ over BEC(ϵ=0.5). The codes with $d_r \leq 14$ satisfy the condition of Corollary 1, while the codes with $d_r > 14$ do not. The fastest approaching speed is attained at $d_r = 14$.

is an extension to BMS channels and a proof for capacity-achievability.

Acknowledgements

The second author would like to thank V. Aref for helping and discussing this work. The second author started this work with V. Aref when he stayed at EPFL in 2011.

References

[1] A. J. Felström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2181–2191, June 1999.

[2] M. Lentmaier, D. V. Truhachev, and K. S. Zigangirov, "To the theory of low-density convolutional codes. II," *Probl. Inf. Transm.* , no. 4, pp. 288–306, 2001.

[3] S. Kudekar, T. Richardson, and R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.

[4] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially Coupled Ensembles Universally Achieve Capacity under Belief Propagation," *ArXiv e-prints*, Jan. 2012.

[5] M. Luby, "LT codes," in *Proc. 40th Annual Allerton Conf. on Commun., Control and Computing*, 2002, pp. 271 – 280.

[6] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, June 2006.

[7] K. Kasai and K. Sakaniwa, "Spatially-coupled MacKay-Neal codes and Hsu-Anastasopoulos codes," *IEICE Trans. Fundamentals*, vol. E94-A, no. 11, pp. 2161–2168, Nov. 2011.

[8] N. Obata, Y.-Y. Jian, K. Kasai, and H. D. Pfister, "Spatially-coupled multi-edge type ldpc codes with bounded degrees that achieve capacity on the bec under bp decoding," Jan. 2013, submitted to ISIT2013.

[9] D. G. M. Mitchell, K. Kasai, M. Lentmaier, and D. J. Costello, "Asymptotic analysis of spatially coupled MacKay-Neal and Hsu-Anastasopoulos LDPC codes," in *2012 International Symposium on Information Theory and its Applications (ISITA)*, Oct. 2012, pp. 337–341.

[10] V. Aref and R. Urbanke, "Universal rateless codes from coupled LT codes," in *Proc. 2011 IEEE Information Thoery Workshop (ITW)*, Oct. 2011.

[11] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

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

[13] S. Pillai, T. Suel, and S. Cha, "The Perron-Frobenius theorem: some of its applications," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 62 – 75, Mar. 2005.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

Appendix

Definition 2: A square matrix A is said to be a reducible matrix when there exists a permutation matrix P such that $P^\mathsf{T} A P = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$, where X and Z are both square. Otherwise A is said to be irreducible.

Definition 3: Let A be a square matrix of size m. The graph $\mathcal{G}(A)$ of A is defined to be the directed graph on m nodes $N_1, ..., N_m$ in which there is a directed edge leading from N_i to N_j if and only if $a_{i,j} \neq 0$. $\mathcal{G}(A)$ is called strongly connected if for each pair of nodes (N_i, N_j) there is a sequence of directed edges leading from N_i to N_j.

The following lemma can be found in [14, p. 362].

Lemma 3: A square matrix A is an irreducible matrix if and only if $\mathcal{G}(A)$ is strongly connected.

Definition 4: We say that a square real matrix $A = (a_{i,j})$ is positive band matrix of width w if

$$a_{i,i+j} \begin{cases} > 0 & (|j| \leq w) \\ = 0 & (|j| > w) \end{cases}$$

Lemma 4: $L \times L$ matrix A_L is irreducible if A_L is positive band matrix of width $w \geq 1$.

Proof: From Definition 4, it holds that for any $0 \leq j \leq L$, $N_j \in \mathcal{G}(A)$ is adjacent to $N_{\max(0,j-1)}$ and $N_{\min(j+1,L-1)}$. The lemma follows readily from Lemma 3. □