Spatially-Coupled MacKay-Neal Codes with No Bit Nodes of Degree Two Achieve the Capacity of BEC

Takuya Okazaki and Kenta Kasai
Department of Communications and Computer Engineering,
Tokyo Institute of Technology
Email: {osa.kenta}@comm.ce.titech.ac.jp

Abstract—Obata et al. proved that spatially-coupled (SC) MacKay-Neal (MN) codes achieve the capacity of BEC. However, the SC-MN codes codes have many variable nodes of degree two and have higher error floors. In this paper, we prove that SC-MN codes with no variable nodes of degree two achieve the capacity of BEC.

I. INTRODUCTION

Felström and Zigangirov introduced spatially-coupled (SC) codes defined by sparse parity check matrix. SC codes are based on construction method for convolutional LDPC codes [1]. Lantuera et al. confirmed that regular SC LDPC codes achieve MAP threshold of original LDPC block codes by BP decoding in at least certain accuracy [2]. Kudekar et al. proved that SC codes achieve MAP threshold by BP decoding on binary erasure channel (BEC) [3] and binary symmetric channel [4]. Kasai et al. introduced SC MacKay-Neal (MN) codes, and showed that these codes with finite maximum degree achieve capacity of BEC by numerical experiment [5]. Obata et al. proved that SC-MN codes achieve capacity of BEC by numerical experiment [6]. It has been observed that SC-MN codes have many variable nodes of degree two. This leads to high error floors.

In this paper, we deal with SC-MN codes whose bit node degree is greater than two. We prove the codes achieve the capacity of BEC. The codes achieve Shannon limit $\epsilon^{\text{Shan}} = 1 - \frac{1}{l}$ for any $l \geq 3$.

II. BACKGROUND

A. MacKay-Neal Codes

($l, r, g$) MN codes are multi-edge type (MET) LDPC codes defined by pair of multi-variables degree distributions ($\mu, \nu$) listed below.

$$\mu(x; \epsilon) = \frac{1}{l} x_1^{l} + \epsilon x_2^g,$$

$$\nu(x) = x_1^{l} x_2^g.$$

In general, the recursion of density evolution of MET-LDPC codes on BEC is given by

$$y_j^{(t)} = 1 - \frac{\mu_j(1-x_j^{(t)};1-\epsilon)}{\mu_j(1;1)}, \quad x_j^{(t+1)} = \frac{\nu_j(y_j^{(t)};\epsilon)}{\nu_j(1;1)},$$

where $x_j^{(t)}$ is probability of erasure message sent along edges of type $j$ at the $t$-th decoding round. Therefore, density evolution of ($l, r, g$) MN codes is

$$x_j^{(t+1)} = f\left(g(x_j^{(t)});\epsilon\right), \quad f(x; \epsilon) = (x_1^{l-1}, x_2^{2g-1}),$$

$$g(x) = (1 - (1-\epsilon)^{-1}(1-x_2^r), 1 - (1-x_1^r)(1-x_2^g)).$$

B. Spatially-Coupled MacKay-Neal Codes

SC-MN codes of coupling number $L$ and of coupling width $w$ are defined by the Tanner graph constructed by the following process. First, at each section $i \in Z$, place $rM/l$ bit nodes of type 1 and $M$ nodes of type 2. Bit nodes of type 1 and 2 are of degree $l$ and $g$, respectively. Next, at each section $i \in Z$, place $M$ check nodes of degree $r + g$. Then, connect edges uniformly at random so that bit nodes of type 1 at $i$ are connected with check nodes at each section $i \in [i, \ldots, i + w - 1]$ with $rM/w$ edges, and bit nodes of type 2 at section $i$ are connected with check nodes at each section $i \in [i, \ldots, i + w - 1]$ with $yM/w$ edges. Bits at section $i \notin [0, L - 1]$ are shortened. Bits of type 1 and 2 at section $i \in [0, L - 1]$ are punctured and transmitted, respectively. Rate of SC-MN codes $R_{MN}$ is given by

$$R_{MN} = \frac{r}{l} + \frac{1 + w - 2 \sum_{i=0}^{w-1}(1 - (\frac{r}{l})^{i+y})}{L} = \frac{r}{l} (L \to \infty).$$

C. Vector Admissible System and Potential Function

In this section, we define vector admissible systems and potential functions.

Definition 1. Define $\mathcal{X} \triangleq [0, 1]^d$, and $F : \mathcal{X} \times [0, 1] \to \mathbb{R}$ and $G : \mathcal{X} \to \mathbb{R}$ as functionals satisfying $G(0) = 0$. Let $D$ be a $d \times d$ positive diagonal matrix. Consider a general recursion defined by

$$x_j^{(t+1)} = f\left(g(x_j^{(t)});\epsilon\right),$$

where $f : \mathcal{X} \times [0, 1] \to \mathcal{X}$ and $g : \mathcal{X} \to \mathcal{X}$ are defined by $F'(x; \epsilon) = f(x; \epsilon)D$ and $G'(x) = g(x)D$, where $F'(x; \epsilon) \triangleq \frac{\partial F(x)}{\partial x_1^{d_1}}, \ldots, \frac{\partial F(x)}{\partial x_d}$. Then the pair $(f, g)$ defines a vector admissible system if

1. $f, g$ are twice continuously differentiable,
2. $f(x; \epsilon)$ and $g(x)$ are non-decreasing in $x$ and $\epsilon$ with respect to $\leq 0$.

1We say $x \leq y$ if $x_i \leq y_i$ for all $1 \leq i \leq d$.
3. \( f(g(0); \epsilon) = 0 \) and \( F(g(0); \epsilon) = 0 \). We say \( x \) is a fixed point if \( x = f(g(x); \epsilon) \).

It can be seen that the density evolution \((f, g)\) of \((l, r, g)\) MN codes given in [7] is a vector admissible system by choosing \( F(x; \epsilon), G(x) \) and \( D \) as below, since this system \((f, g)\) satisfies the condition in Definition 1

\[
F(x; \epsilon) = \frac{r}{n} + \epsilon x_2^2, \\
G(x) = r x_1 + g x_2 + (1 - x_1)^r (1 - x_2)^g - 1, \\
D = \begin{pmatrix} r & 0 \\ 0 & g \end{pmatrix}.
\]

Definition 2 ([7] Def. 2). We define the potential function \( U(x; \epsilon) \) of a vector admissible system \((f, g)\) by

\[
U(x; \epsilon) \triangleq g(x) D x^T - G(x) - F(g(x); \epsilon).
\]

The potential function \( U(x_1, x_2, \epsilon) \) of \((l, r, g)\) MN codes is given by

\[
U(x_1, x_2, \epsilon) = 1 - \epsilon \left( (1 - (1 - x_1)^r) (1 - x_2)^g - 1 \right) - r (1 - (1 - x_1)^r - 1 (1-x_2)^g) + (1 - x_1)^r (1 - x_2)^g \left( 1 + \frac{r(x_1)}{1-x_1} + \frac{g(x_2)}{1-x_2} \right).
\]

Definition 3 ([7] Def. 7). Let \( \mathcal{F}(\epsilon) \triangleq \{ x \in X \{0\} \mid x = f(g(x); \epsilon) \} \) be a set of non-zero fixed points for \( \epsilon \in [0, 1] \). The potential threshold \( \epsilon^* \) is defined by

\[
\epsilon^* \triangleq \sup \{ \epsilon \in [0, 1] \mid \min \mathcal{F}(\epsilon) U(x; \epsilon) > 0 \}.
\]

Let \( \epsilon^*_s \) be threshold of uncoupled system defined in [7] Def. 6]. For \( \epsilon \) such that \( \epsilon^*_s < \epsilon < \epsilon^* \), we define energy gap \( \Delta E(\epsilon) \) as

\[
\Delta E(\epsilon) \triangleq \max_{\epsilon' \in [\epsilon, 1]} \inf_{x \in \mathcal{F}(\epsilon')} U(x; \epsilon').
\]

We define the SC system of a vector admissible system.

Definition 4 ([7] Def. 9). For a vector admissible system \((f, g)\), we define the SC system of coupling number \( L \) and coupling width \( w \) as

\[
x_i^{(t+1)} = \frac{1}{w} \sum_{k=0}^{w-1} f \left( \frac{1}{w} \sum_{j=0}^{w-1} g(x_{i+j-k}^{(t)}); \epsilon_{i-k} \right), \\
\epsilon_i = \begin{cases} \epsilon, & i \in \{0, \ldots, L-1\} \\ 0, & i \notin \{0, \ldots, L-1\} \end{cases}
\]

If we define \((f, g)\) as the density evolution for \((l, r, g)\) MN codes in [7], the SC system gives the density evolution of SC-MN codes with coupling number \( L \) and coupling width \( w \).

Next theorem states that if \( \epsilon < \epsilon^* \) then fixed points of SC vector system converge towards 0 for sufficiently large \( w \).

Theorem 1 ([7] Thm. 1]). Consider the constant \( K_{f, g} \) defined in [7] Lem. 11. This constant value depends only on \((f, g)\). If \( \epsilon < \epsilon^* \) and \( w > (dK_{f, g})/(2\Delta E(\epsilon)) \), then the SC system of \((f, g)\) with coupling number \( L \) and coupling width \( w \) has a unique fixed point 0.

We will show that the potential threshold \( \epsilon^* \) of \((l, r = 3, g = 3)\) MN codes is \( 1 - R^MN = 1 - 3/l \) for any \( l \geq 3 \). This is sufficient to show that \((l, 3, 3)\) SC-MN codes with sufficiently large \( w \) and \( L \) achieve the capacity of BEC under BP decoding.

III. PROOF OF ACHIEVING CAPACITY

In this section, we calculate the potential threshold \( \epsilon^* \) of \((l, r = 3, g = 3)\) MN codes. To this end, we first investigate the set of fixed points \( \mathcal{F}(\epsilon) \).

The density evolution recursion in [7] can be rewritten as

\[
x_1^{(t+1)} = (1 - (1 - x_1)^r) - 1 (1 - x_2)^g - 1, \\
x_2^{(t+1)} = \epsilon (1 - (1 - x_1)^r (1 - x_2)^g - 1).
\]

Fixed points \((x_1, x_2; \epsilon)\) of density evolution with \( x_1 = 0 \) and \( x_1 = 1 \) are \((0, 0; \epsilon)\) and \((1, 1; \epsilon)\), respectively. We define these fixed points as trivial fixed points and all other fixed points as non-trivial fixed points. All non-trivial fixed points \((x_1, x_2; \epsilon)\) can be parametrically described as

\[
x_2(x_1) = 1 - \left( \frac{1 - x_1}{1 - x_1} \right)^{1/r}, \\
\epsilon(x_1) = \frac{x_2(x_1)}{1 - (1 - x_1)^r (1 - x_2(x_1))^{g-1})^{g-1}}.
\]

with \( x_1 \in (0, 1) \).

Next, we shall investigate the value of the potential function at the fixed points. The value of the potential functions at trivial fixed point \((1, \epsilon; \epsilon)\) is respectively given by

\[
U(1, \epsilon, \epsilon) = 1 - \frac{r}{l} - \epsilon.
\]

Figure 1 draws the potential function of \((l, r, g)\) MN codes at fixed points \( \epsilon \in \mathcal{F}(\epsilon) \). It appears that the potential function at non-trivial fixed points is always positive. We will prove this.
To be precise, the potential function of \((l, r, g)\) MN codes for non-trivial fixed points satisfies
\[
U(x_1, x_2(x_1), \epsilon(x_1)) > 0 \quad \text{for} \quad x_1 \in (0, 1).
\] (4)

Our strategy of proof is as follows. First change the representation of (4) into a polynomial form by changing variables a few times. Then apply Sturm’s theorem for smaller \(l\) and bound the polynomial for larger \(l\).

We define \(U(z) := U(x_1, x_2(x_1), \epsilon(x_1))|_{x_1=2^{-l-1}}\). Obviously, to prove (4), it is sufficient to show \(U(z) > 0\) for \(z \in (0, 1)\).

\[
U(z) = \frac{-3z^l}{l} + (1 - z)(1 - 4z^{l-1}) + (1 - z)^{1/3}(1 - z^{l-1})^{-2/3} - 2(1 - z)^{2/3}(1 - z^{l-1})^{5/3}.
\]

We use next lemma to eliminate fractional power in \(U(z)\). The proof is given in Section IV-A.

**Lemma 1.** Define \(H(u, z)\) as follows.

\[
H(u, z) = (u + \frac{3z^l}{l} - (1 - z)(1 - 4z^{l-1}))^3 + 6(1 - z)(1 - z^{l-1})(u + \frac{3z^l}{l} - (1 - z)(1 - 4z^{l-1})) - (1 - z)(1 - z^{l-1})^{-2} + 8(1 - z)^2(1 - z^{l-1})^5.
\]

Then, \(H(0, z) < 0\) for \(z \in (0, 1)\) implies \(U(z) > 0\) for \(z \in (0, 1)\).

Define \(I(z) := \frac{l^3(1 - z^{l-1})^2}{(1 - z)^2}H(0, z)\). Obviously, to prove \(H(0, z) < 0\) for \(z \in (0, 1)\), it is sufficient to prove \(I(z) < 0\) for \(z \in (0, 1)\). We see that \(I(z)\) for \(l \geq 3\) is a polynomial as follows.

\[
I(z) = -l^3 + 27\sum_{i=0}^{l-2} \left[3^{i-2}2^i(1 - z^{l-1})ight]
- 27l^22^i(1 - 4z^{l-1})(1 - z^{l-1})^2
- 9l^2(1 - z^{l-1})^2\{(3 + z)z^2
+ 16(-1 + z)z^{2i} - 8(-1 + z)z^{1+i}\}
- l^3(1 - z)^{-9+i}\left\{8z^{6i} - 56z^{3+5i} + 2z^6(3 + 7z)
+ 8z^{2+4i}(13 + 8z) - 8z^{3+3i}(13 + 22z)
+ 4z^{4+2i}(21 + 43z) - z^{5+i}(41 + 73z)\right\}.
\] (5)

We prove \(I(z) < 0\) for \(3 \leq l < 166\) and \(l \geq 166\) in the following lemmas. The proofs are given in Section IV-B and Section IV-C respectively.

**Lemma 2.** For \(3 \leq l < 166\), \(I(z) < 0\) for \(z \in (0, 1)\).

**Lemma 3.** For \(l \geq 166\), \(I(z) < 0\) for \(z \in (0, 1)\).

**Theorem 2.** For any \(l \geq 3\) and \(\epsilon < e^{\text{Sh}a} = 1 - \frac{3}{4}\), the unique fixed point of density evolution of \((l, 3, 3)\) SC-MN codes of coupling number \(L\) and coupling width \(w\) is \(0\) for sufficiently large \(w\) and \(L\).

Proof: From (4), potential function for non-trivial fixed points is always positive. Therefore, from Definition \ref{def} and potential function for trivial fixed point \(\epsilon = e^{\text{Sh}a}\). From Theorem \ref{thm} for \(\epsilon < e^{\text{Sh}a}\), the unique fixed point of density evolution for \((l, 3, 3)\) SC-MN codes is \(0\).

The case with \(l = 3\) implies rate one codes over BEC(0). Some might think this is not interesting. Nevertheless, we included the case with \(l = 3\) for comprehensiveness.

**IV. PROOF OF LEMMAS**

**A. Proof of Lemma \ref{lemma}**

Partial derivative of \(H(u, z)\) with respect to \(u\) gives

\[
\frac{\partial H(u, z)}{\partial u} = 3(u + \frac{3z^l}{l} - (1 - z)(1 - 4z^{l-1}))^2 + 6(1 - z)(1 - z^{l-1}) \geq 0.
\] (6)

Substituting \(u = U(z)\) into \(H(u, z)\) gives

\[
H(U(z), z) = ((1 - z)^{1/3}(1 - z^{l-1})^{-2/3} - 2(1 - z)^{2/3}(1 - z^{l-1})^{5/3})^3
+ 6(1 - z)(1 - z^{l-1})\{(1 - z)^{1/3}(1 - z^{l-1})^{-2/3} - 2(1 - z)^{2/3}(1 - z^{l-1})^{5/3}\}
- (1 - z)(1 - z^{l-1})^{-2} + 8(1 - z)^2(1 - z^{l-1})^5
= (1 - z)(1 - z^{l-1})^{-2} - 8(1 - z)^2(1 - z^{l-1})^5
- 6(1 - z)(1 - z^{l-1})\{(1 - z)^{1/3}(1 - z^{l-1})^{-2/3} - 2(1 - z)^{2/3}(1 - z^{l-1})^{5/3}\}
+ 6(1 - z)(1 - z^{l-1})\{(1 - z)^{1/3}(1 - z^{l-1})^{-2/3} - 2(1 - z)^{2/3}(1 - z^{l-1})^{5/3}\}
- (1 - z)(1 - z^{l-1})^{-2} + 8(1 - z)^2(1 - z^{l-1})^5
= 0.
\] (7)

From (6), \(H(u, z)\) monotonically increasing with respect to \(u\). From (7), \((u, z) = (U(z), z)\) is a root of \(H(u, z) = 0\). Therefore \(H(0, z) < 0\) for \(z \in (0, 1)\) implies \(U(z) > 0\) for \(z \in (0, 1)\).

**B. Proof of Lemma \ref{lemma2}**

From \(I(0) = -l^3\) and \(I(1) = -l^3\), we see that \(z = 0, 1\) are not multiple roots of equation \(I(z) = 0\). Let \(I_1(z), \ldots, I_m(z)\) be Sturm sequences of \(I(x)\). Let \(V(z)\) be the number of sign changes in the sequence. Table \ref{tab1} lists sign changes of Sturm sequence \(I_1(z), \ldots, I_m(z)\) of \(I(x)\) in (5) for \(l = 3, \ldots, 11\). \(V(z)\) is the number of sign changes in the sequence. We see that \(V(0) = V(1)\). We observed that \(V(0) = V(1)\) for \(l < 165\) but not listed all due to the space limit. From Theorem \ref{thm2} this implies that the number of distinct roots of equation \(I(z) = 0\) in \((0, 1)\) is \(V(0) - V(1) = 0\). Therefore, \(I(z) < 0\), \(z \in (0, 1)\) for \(3, \ldots, 164\).
Since \( g \) gives the maximum value of \((b), we apply (8) and (9) to each term of (5) by using \( d \). Differentiating \( d \) gives
\[
q(z) := z^{a+b}(1 - z) \leq \frac{1}{a + b + 1}.
\]
If \( \frac{a + b + 1}{a + b + 1} \in (0, 1) \), then
\[
q(z) := z^{a+b}(1 - z) \leq \frac{1}{a + b + 1}.
\]
Since \( \frac{a + b + 1}{a + b + 1} \in (0, 1) \), we see that \( z = \frac{a + b}{a + b + 1} \) gives the maximum value of \( q(z) \).
\[
q(z) \leq \left( \frac{a + b}{a + b + 1} \right)^{a+b} = \left( \frac{a + b}{a + b + 1} \right)^{a+b} \frac{1}{a + b + 1} < \frac{1}{a + b + 1}.
\]
Differentiating \( r(z) \) gives
\[
\frac{dr(z)}{dz} = z^{a+b-1}(1 - z^{-l-1})((a + b) - ((a + 2)l + b - 2)z^{l-1}).
\]
We obtain an upper bound of \( I(z) \) for \( z \in (0, 1) \) as follows.
\[
I(z) < -l^3 + 27 \left( \sum_{i=0}^{l-2} z^{3l-2+i}(1 - z^{l-1}) + 108l^2z^{-3+3l}(1 - z^{l-1})^2 + 9l^3(3z^{2l - 2} + 8z^{2l + 1}) + l^3(5z^{5+1}(13 + 22z)) \right)
\]
\[
< -l^3 + 27 \left( \sum_{i=0}^{l-2} [1 + 108l^2(\frac{2l - 2}{5l - 5})^2 + 9l^3(\frac{2l - 2}{3l - 4})^2 + 8(\frac{2l - 2}{4l - 4})^2] + l^3(\frac{56}{6l - 7} + 176\frac{4l - 5}{4l - 5} + 104\frac{2l - 3}{2l - 3} + 73\frac{73}{2l - 2}) \right)
\]
\[
< -l^3 + 27(l - 1) + 432l^2 + 9l^3(\frac{5}{9} + 16\frac{1}{9} + 8\frac{1}{4}) + 5l^3(\frac{59}{29l} + 176\frac{1}{19l} + 104\frac{73}{9f} + 41\frac{9f}{19f})
\]
\[
< -l^3 + \frac{677534}{41325}l^2 + 444\frac{1}{5}l =: T(l).
\]
In (a), we eliminate negative terms except for \(-l^3\). Next, in (b), we apply (3) and (2) to each term of (5) by using \( l \geq 165 \).
V. Conclusion and Future Work

In this paper, we proved that \((l, 3, 3)\) SC-MN codes with \(l \geq 3\) achieve capacity on the BEC under BP decoding for sufficiently large \(L\) and \(w\). This codes do not have bit nodes of degree two and have low error floors. We proved that the potential threshold and Shannon limit of \((l, r = 3, g = 3)\) MN codes on the BEC are the same.

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] ——, “Spatially coupled ensembles universally achieve capacity under belief propagation,” IEEE Trans. Inf. Theory, vol. 59, no. 12, pp. 7761–7813, 2013.
[5] 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.
[6] 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,” in Proc. 2013 IEEE Int. Symp. Inf. Theory (ISIT), July 2013, pp. 2433–2437.
[7] A. Yedla, Y.-Y. Jian, P. Nguyen, and H. Pfister, “A simple proof of threshold saturation for coupled scalar recursions,” in Turbo Codes and Iterative Information Processing (ISTC), 2012 7th International Symposium on, 2012, pp. 51–55.
[8] W. Gautschi, Numerical Analysis, ser. SpringerLink : Bücher. Springer Science+Business Media, LLC, 2011.

Appendix

Sturm’s Theorem

**Theorem 3 (8).** For a polynomial \(f(x)\) over \(\mathbb{R}\), we define Sturm sequences \(f_i(x) (i = 0, \ldots, m)\) as \(f(x)\), \(f'(x)\) and polynomials obtained by applying Euclid’s algorithm to \(f(x)\) and \(f'(x)\).

\[
egin{align*}
    f_0(x) &= f(x), \\
    f_1(x) &= f'(x), \\
    f_{n-1}(x) &= q_n(x)f_n(x) - f_{n+1}(x) \quad (n = 1, \ldots, m - 1), \\
    f_{m-1}(x) &= q_m(x)f_m(x).
\end{align*}
\]

For real number \(c\), let \(V(c)\) be the number of sign changes in \(f_0(c), f_1(c), \ldots, f_m(c)\). If neither \(a \in \mathbb{R}\) nor \(b \in \mathbb{R}\) is a multiple root of \(f(x) = 0\), then the number of distinct roots of \(f(x)\) in \((a, b)\) is \(V(a) - V(b)\).