在外研究員研究報告書

２０１８年３月23 日　受付

所　属	理工学部	氏　名	程　俊　（程）
職　名	教授		
研究課題名	通信路符号化技術および信号処理技術に関する基礎研究		
研究期間	２０１６年１０月１日　～２０１７年９月３０日		

滞在期間・滞在地研究調査先	滞在期間	滞在地	研究・調査先
	2016/10/1-2017/9/30	College Station, Texas, USA	Texas A&M University

研究費	306.6 万円	研究成果の概要	別記　4,000字程度

発表	題目名	発表学術誌名Vol. No.	発行年月日
	Signature design of sparsely spread code division multiple access based on superposed constellation distance analysis	IEEE Access, vol. 5, no. 1, pp. 23809-23821	Oct. 2017
	題目名	**発表学術誌名Vol. No.**	**発行年月日**
	Improved sphere bound on the MLD performance of binary linear block codes via Voronoi region	IEICE Trans. Fundamentals, vol. E100-A, no. 12, pp. 2572-2577	Dec. 2017
	演題	**講演学会名**	**講演年月日**
	A user-independent serial interference cancellation based coding scheme for the unsourced random access Gaussian channel	IEEE Information Theory Workshop	Nov. 6, 2017

A Coding Scheme for Unsourced Random Access Gaussian Channel[§]

Jun Cheng
Department of Intelligent Information Engineering and Sciences
Doshisha University

Abstract

A coding scheme for the unsourced multiple access channel is proposed. This new paradigm is composed of four main ingredients: (i) the transmission period is partitioned into sub-blocks, thereby instituting a slotted framework; (ii) The message (data) is split into two parts and one part chooses an interleaver for a low density parity check (LDPC) type code. This part of the message is encoded using spreading sequences or codewords that are designed to be decoded by a compressed sensing type decoder; (iii) The other part of the message is encoded using a LDPC type code and decoded using a joint message passing decoding algorithm designed for the T-user binary input real adder channel; (iv) users repeat their codeword in multiple sub-blocks, with the transmission pattern being a deterministic function of message content and independent of the identity of the user. When this coding scheme is combined with serial interference cancellation, the ensuing communication infrastructure can offer significant performance improvements compared to the recently proposed coding scheme in [2] and results in the best performing coding scheme to date.

I. Problem Description

In [1], Polyanskiy introduced an interesting and timely multiple access problem*; throughout, we refer to this new formulation as the unsourced multiple access channel model (MAC). In this setting, a very large number, K_{tot}, of users in a wireless network operate in an uncoordinated fashion. Out of the K_{tot} users, a subset of K_{a} users are active at any time; and each of them wishes to communicate a B-bit message to a central base station. The base station is interested only in recovering the list of messages without regard to the identity of the user who transmitted a particular message. In addition to this, the interest is typically in the case when B is small.

Specifically, let K_{tot} and K_{a} denote the total number of users in the network and the number of active users, respectively. Each user has B bits of information (or, one of $M := 2^B$ indices) to be encoded and transmitted within a time frame of $\tilde{N}$ uses of the channel. In addition to these, various important parameters encountered in the report along with their notation are summarized in Table. I. Let $W_i \in [1 : M]$ be a random variable that represents the message index of the i-th user and let w_i be a realization of the random variable W_i. Throughout this report $[a : b]$ denotes the set of integers from a to b, both inclusive. We first assume that W_i is uniformly distributed over the set $[1 : M]$ and that for any pair of users i and j, W_i and W_j are independent of each other. Without loss of generality we assume that the set of indices of the active users is $[1 : K_{\text{a}}]$ and let $\mathcal{W} := \{w_1, w_2, \ldots, w_{K_{\text{a}}}\}$ denote the set of transmitted messages.

The observed signal vector at the receiver corresponding to the $\tilde{N}$ channel uses can be written as

$$\vec{y} = \sum_{i=1}^{K_{\text{tot}}} \vec{x}_i + \vec{z}, \tag{1}$$

[§]This is a collaborative work with Dr. Avinash Vem, Profs. Krishna Narayanan, and Jean-Francois Chamberland, Department of Electrical and Computer Engineering, Texas A&M University.

*I knew the problem in Information Theory and Applications Workshop, hold Feb. 12-17, 2017, in Catamaran Resort, Pacific Beach, San Diego. In the workshop, Dr. Or Ordentlich, from MIT, gave an oral presentation titled "low complexity schemes for the random access Gaussian MAC," and reported a low complexity solution to this problem. Motivated by Drs. Ordentlich and Polyanskiy's pioneering work, I had an idea to attempt to solve the problem. I discussed with Prof. Krishna Narayanan during the workshop, and started this work. The contribution of this work is summarized in this report.

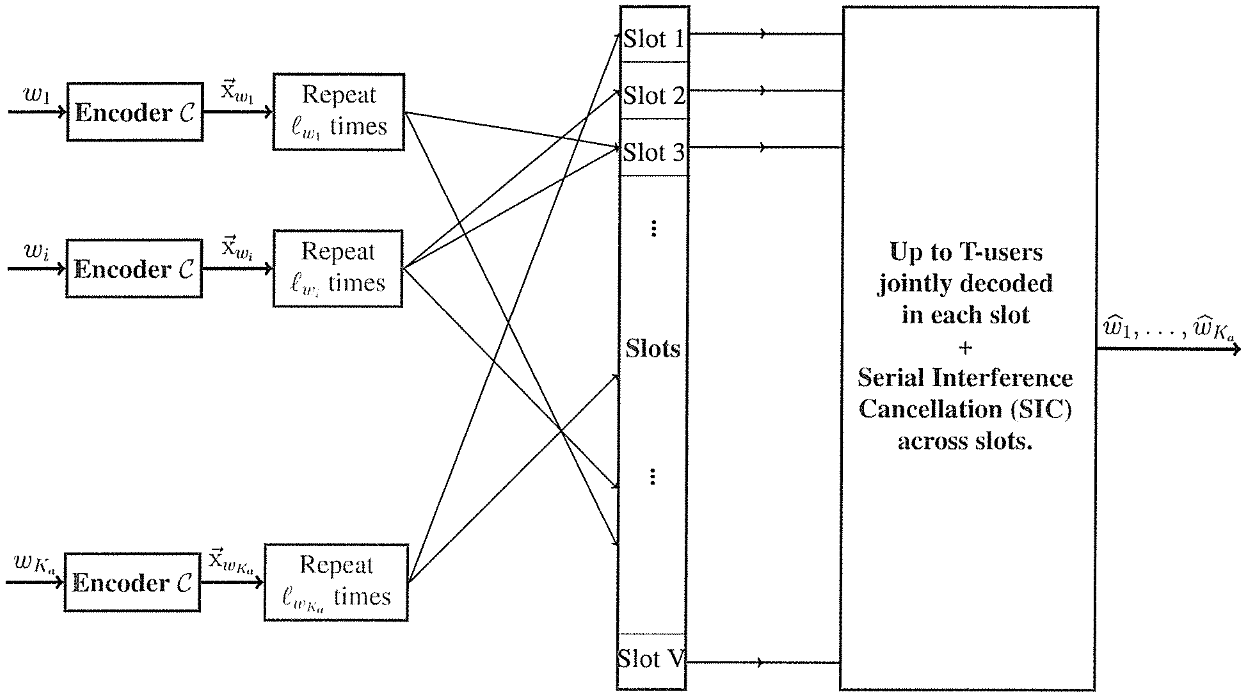

Fig. 1: Schematic of the proposed scheme

B. Transmission policy within a sub-block - same code book scheme for the T-user multiple access

There are two components to the code $\mathcal{C}$ used in the proposed transmission scheme within each sub-block: a good sensing matrix for a T-sparse robust compressed sensing (CS) problem and a good channel code for the T-user binary-input real-adder channel that is decodable with low computational complexity. The B bits to be transmitted are split into two groups of size B_p and $B_\text{c} = B - B_\text{p}$ bits, respectively. For convenience, we define $M_\text{p} := 2^{B_\text{p}}$ and $M_\text{c} := 2^{B_\text{c}}$.

The description of the overall encoder for code book $\mathcal{C}$ combining the above two components can be described as following. Let $w = (w^\text{p}, w^\text{c})$ be the message index to be encoded, where the indices w^p and w^c correspond to the preamble and coding message indices respectively. We first encode the message index w^c to the codeword $\vec{c}_{w^\text{c}} \in \mathcal{C}_\text{c}$ followed by permuting it according to permutation $\pi_{\tau_{w^\text{p}}} = [\pi^1_{\tau_{w^\text{p}}}, \pi^2_{\tau_{w^\text{p}}}, \dots, \pi^{N_\text{c}}_{\tau_{w^\text{p}}}]$. The final code word $\vec{c}_w$ is then obtained by inserting the w^pth column from the compressed sensing matrix $\mathbf{A}$ at the beginning of the permuted codeword i.e.,

$$\begin{aligned}\vec{c}_w &= [\vec{a}_{w^\text{p}}, \pi_{\tau_{w^\text{p}}}(\vec{c}_{w^\text{c}})] \qquad \text{where } \vec{a}_{w^\text{p}} \in \mathbf{A}, \vec{c}_{w^\text{c}} \in \mathcal{C}_\text{c} \\ &= [\vec{a}_{w^\text{p}}, c_{w^\text{c}}(\pi^1_{\tau_{w^\text{p}}}), c_{w^\text{c}}(\pi^2_{\tau_{w^\text{p}}}), \dots, c_{w^\text{c}}(\pi^{N_\text{c}}_{\tau_{w^\text{p}}})].\end{aligned} \tag{4}$$

The overall encoding process is summarized in Fig. 2.

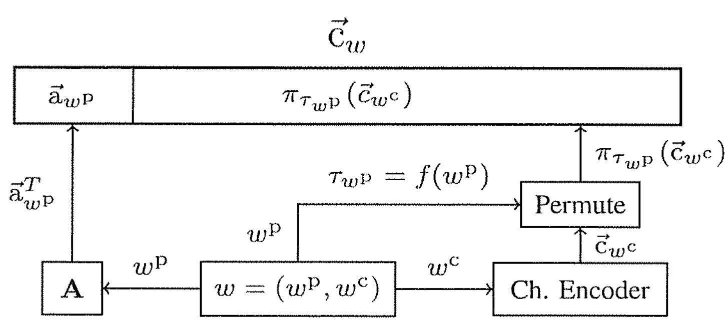

Fig. 2: Schematic depicting the overall encoding scheme in a sub-block given the message index $w = (w^\text{p}, w^\text{c})$. The final code word transmitted in a sub-block is given by $\vec{c}_w = [\vec{a}_{w^\text{p}}, \pi_{\tau_{w^\text{p}}}(\vec{c}_{w^\text{c}})]$.

coding scheme for the unsourced GMAC channel the authors Ordentlich and Polyanskiy also evaluate the performance of their proposed scheme by computing the minimum SNR required to achieve the target error probability for a fixed set of parameters. To make the comparison convenient we pick identical parameters, summarized as following:

- number of bits each user intends to transmit $B = 100$
- total number of channel uses $\tilde{N} = 30,000$
- number of active users $K_a \in [25 : 300]$
- maximum per user error probability $P_e \leq \varepsilon = 0.05$.

With the parameters $B, \tilde{N}, K_a, \varepsilon$ fixed, the choices for the design parameters are as following:

1) Maximum number of users to be jointly decoded at a slot $T \in \{2, 4, 5\}$.
2) The left d.d is chosen to be $L(x) = \beta x + (1 - \beta)x^2$. The free parameter is optimized over the set $\beta \in \{0, 0.1, \dots, 1\}$.
3) Number of preamble and channel coding message bits: $B_p = 9, \ B_c = B - B_p = 91$.
4) *Sensing matrix for preamble component*: Note that $M_p = 2^{B_p} = 512$ is the size of the sensing matrix $\mathbf{A}$. We choose the sensing matrix of dimensions $N_p \times M_p = 63 \times 512$.
5) *Channel coding component*: The number of channel uses available for channel coding N_c is dependent on N which in turn depends on the total number of sub-blocks V. It is impractical to build a channel code for various rates $R_c = \frac{B_c}{N_c}$ (although B_c is fixed, N_c needs to be optimized over) and evaluate the performance numerically for each set of parameters (N_c, B_c). Therefore to evaluate the performance of the channel coding component we use the finite block length achievability bound due to Polyanskiy. This seems a reasonable choice that one can construct LDPC codes even for moderate block lengths that perform close to the above mentioned bound.

For performace evaluation, at j-th sub-block where $R_j \leq T$, let us define the following error events:

- $\mathcal{E}_{pj}$: Given there is no preamble collision, let $\mathcal{E}_{pj}$ be the event that the output of the compressed sensing decoder is incorrect i.e., $\widehat{\mathcal{W}_j^p} \neq \mathcal{W}_j^p$. The event $\mathcal{E}_p$ is defined for the worst case $R_j = T$
- $\mathcal{E}_{ej}$: Let $\mathcal{E}_{ej}$ be the event that the error energy test makes an error. With the following notations:
  - Given that there is no preamble collision and the compressed sensing decoder is correct let $\mathcal{E}_{ej}^0$ be the event the error energy test detects a preamble collision and
  - let $\mathcal{E}_{ej}^1$ be the event there exists a collision but the energy test fails to detect the preamble collision

  we can see that $\mathcal{E}_{ej} = \mathcal{E}_{ej}^0 \cup \mathcal{E}_{cj}^1$.
- $\mathcal{E}_{cj}$: Given there is no preamble collision and that the preamble message indices are decoded successfully, let $\mathcal{E}_{cj}$ be the event that the channel decoder fails to recover all the channel coding message indices correctly. The event $\mathcal{E}_c$ is defined for the worst case, when $R_j = T$
- $\mathcal{E}_{\text{SIC}}$: Let $\mathcal{E}_{\text{SIC}}$ be the event that a random user is not recovered by the SIC decoding process

We observe that the overall decoding process within a given sub-block j making an error is a disjoint union of the above described events i.e.,

$$\mathcal{E}_j = \mathcal{E}_{pj} \cup \mathcal{E}_{ej}^0 \cup \mathcal{E}_{cj} \cup \mathcal{E}_{ej}^1 = \mathcal{E}_{pj} \cup \mathcal{E}_{ej} \cup \mathcal{E}_{cj}.$$

The per user error probability P_e, which is equivalent to $\Pr(\mathcal{E}_{\text{SIC}})$, can be bounded as following:

$$\begin{aligned}
P_e = \Pr(\mathcal{E}_{\text{SIC}}) &\leq \Pr\left(\mathcal{E}_{\text{SIC}} | \, (\bigcup_j \mathcal{E}_j)^c\right) + \Pr\left(\bigcup_j \mathcal{E}_j\right) \\
&\leq \Pr\left(\mathcal{E}_{\text{SIC}} | \, \bigcap_j \mathcal{E}_j^c\right) + \Pr\left(\bigcup_j \mathcal{E}_{pj} \cup \mathcal{E}_{ej} \cup \mathcal{E}_{cj}\right) \\
&\leq \Pr\left(\mathcal{E}_{\text{SIC}}^{'}\right) + \sum_j (\Pr(\mathcal{E}_{pj}) + \Pr(\mathcal{E}_{ej}) \Pr(\mathcal{E}_{cj})) \\
&\leq \Pr\left(\mathcal{E}_{\text{SIC}}^{'}\right) + V\left(\Pr(\mathcal{E}_p) + \Pr(\mathcal{E}_e) + \Pr(\mathcal{E}_c)\right)
\end{aligned} \tag{5}$$

The x mark represents our proposed scheme where for the channel coding part instead of the FBL bounds we use the actual simulation results. We use a rate-1/4 $(364, 91)$ LDPC code obtained from repeating every coded bit of (3,6) LDPC code twice and a message passing decoder for $T = 2$. It can be seen that the simulation results with the $(3, 6)$ LDPC code are only 0.5 dB away from the curve corresponding to $T = 2$ showing that the pragmatic coding scheme can perform close to the finite length bounds. It can also be seen that our proposed scheme provides substantial gain over the results in [2].

In the proposed encoding scheme, for $L(x) = \beta x + (1 - \beta)x^2$ each user may transmit once or twice depending on the message index chosen. We need to point out that the power constraint employed is an average over all the message indices i.e., $\mathbb{E}_w [||\vec{c}_w||^2] = (2 - \beta)P$.

IV. Conclusion

We have proposed a novel coding scheme for the unsourced multiple access channel model introduced by Polyanskiy [1]. This new paradigm is composed of four main ingredients: (i) the transmission period is partitioned into sub-blocks, thereby instituting a slotted framework; (ii) The message (data) is split into two parts and one part chooses an interleaver for a low density parity check (LDPC) type code. This part of the message is encoded using spreading sequences or codewords that are designed to be decoded by a compressed sensing type decoder; (iii) The other part of the message is encoded using a low density parity check (LDPC) type code and decoded using a joint message passing decoding algorithm designed for the T-user binary input real adder channel; (iv) users repeat their codeword in multiple sub-blocks, with the transmission pattern being a deterministic function of message content and independent of the identity of the user. When this coding scheme is combined with serial interference cancellation, the ensuing communication infrastructure can offer significant performance improvements compared to the recently proposed coding scheme in [2] and results in the best performing coding scheme to date.

The main contribution of this work is to propose, analyze and optimize a new coding architecture that overcomes these drawbacks and substantially improves performance when compared to the state-of-the-art. Key features of our scheme are summarized as follows.

- **User Symmetry:** Active users employ the same coding scheme, with transmitted signals determined solely by the message to be transmitted and is independent of the identity of the user. To be precise, no parameter of the encoding scheme such as the interleaver and spreading sequence are unique to a transmitter.
- **Binary-input, real-adder channel:** The proposed coding scheme is tailored to the binary-input real-adder channel. The information message is split into two parts. The first portion picks an interleaver for an LDPC code, and the second part is encoded using this LDPC code. Bits associated with the first portion are communicated using a compressed sensing scheme. The second part is decoded using a message passing decoder that jointly recovers up to T messages within a slot.
- **Successive interference cancellation:** Active users repeat their codewords in several slots. The repetition patterns are selected based on message bits. This scheme facilitates interference cancellation within the slotted structure, and therefore renders obsolete the over-provisioning of slots to avoid undue collisions with more than T users.

While [2] also incorporates the user symmetry aspect described above, our scheme differs from theirs in the other features highlighted above.

References

[1] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. Int. Symp. on Information Theory*, pp. 2523–2527, IEEE, 2017.

[2] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. Int. Symp. on Information Theory*, pp. 2528–2532, IEEE, 2017.

[3] A. Vem, K. R. Narayanan, J. Cheng, and J.-F. Chamberland, "A user-independent serial interference cancellation based coding scheme for the unsourced random access Gaussian channel," in *Proc. IEEE Information Theory Workshop*, pp. 1-6, Kaohsiung, Taiwan, 2017.