最新结果 • RECENT RESULTS •

Finite-length Code and Application in Network Coding
有限长度编码及在网络编码中的应用

Shenghao Yang, Yanyan Dong and Jie Wang, The Chinese University of Hong Kong, Shenzhen
杨升浩, 董岩岩, 王捷, 香港中文大学（深圳）
shyang@cuhk.edu.cn

A binary code of block length n and codebook size 2^k is called an (n,k) code, which is said to be *linear* if it is a subspace of $\{0,1\}^n$. Linear codes have been extensively studied in coding theory. For memoryless binary symmetric channels (BSCs), asymptotically capacity achieving linear codes with low encoding/decoding complexity have been designed. However for fixed n and k, whether linear codes are optimal or not among all (n,k) codes for BSCs in terms of the maximum likelihood (ML) decoding is a long-standing open problem, dated back to Slepian's 1956 paper [1]. Except for codes that are perfect or quasi-perfect, very little is known about optimal codes for BSC. The best linear $(n,2)$ codes have been explicitly characterized for each block length n [2][3], but whether linear $(n,2)$ codes are optimal or not among all $(n,2)$ codes in terms of the ML decoding is unknown in general.

Recently we derived a general approach for comparing the ML decoding performance of two $(n,2)$ codes with certain small difference. Based on this approach, we verify that linear $(n,2)$ codes are optimal for a range of n. In particular, we show that for any block-length n, there exists an optimal $(n,2)$ code that is either linear or in a subset of nonlinear codes, called Class-I codes. Based on the analysis of Class-I codes, we derive sufficient conditions such that linear codes are optimal. For $n \leq 8$, our analytical results show that linear codes are optimal. For n up to 300, numerical evaluations show that linear codes are optimal, where the evaluation complexity is $O(n^7)$. Moreover, most ML decoding comparison results obtained are universal in the sense that they do not depend on the crossover probability of the BSC.

Finite-length codes find applications in network coding. When studying the communication through a line network with buffer size constraints at intermediate nodes, a class of batched codes are used [5][6]. A batched code has an outer code and an inner code. The outer code encodes the information messages into batches, each of which is a sequence of coded symbols, while the inner code performs a general network coding for the symbols belonging to the same batch. The inner code, comprising of recoding at network nodes on each batch separately, should be designed for specific channels.

Batched codes provide a general coding framework for line networks with buffer size constraints at the intermediate nodes. We have the following results about the performance of batched codes [5][6][7]: When buffer size is a constant of the network length L, the maximum achievable rate decreases exponentially with the network length. When the batch size is a constant, using a buffer size of $O(\log\log L)$ can achieve rate $\Omega\left(\frac{1}{\log L}\right)$, and the achievable rate is upper bounded by $O\left(\frac{1}{\log L}\right)$ as long as the buffer size is $O(\log L)$. Moreover, using $O(\log L)$ batch size and $O(\log L)$ buffer size, the maximum achievable rate can be arbitrarily close to the cut-set bound.

Finite-length codes can be used to form the inner code of batched codes. For line networks of packet erasure channels, random linear codes can be used between two adjacent nodes [8]. Here the random linear codes are used as erasure codes, but do not need to be decoded in each hop. For line networks of binary symmetric channels, it has been shown that [5] when using batch size 1, the inner code formed by repetition codes, i.e., the optimal $(n, 1)$ codes, can achieve the rate $\Omega\left(\frac{1}{\log L}\right)$, which is optimal in terms of scalability with L. To achieve higher absolute rates, larger batch sizes should be used (see Fig. 1). Note that for any fixed batch size k and network length L, the best rate in the figure is achieved by the inner code formed by (n, k) codes with $n = O(\log L)$.

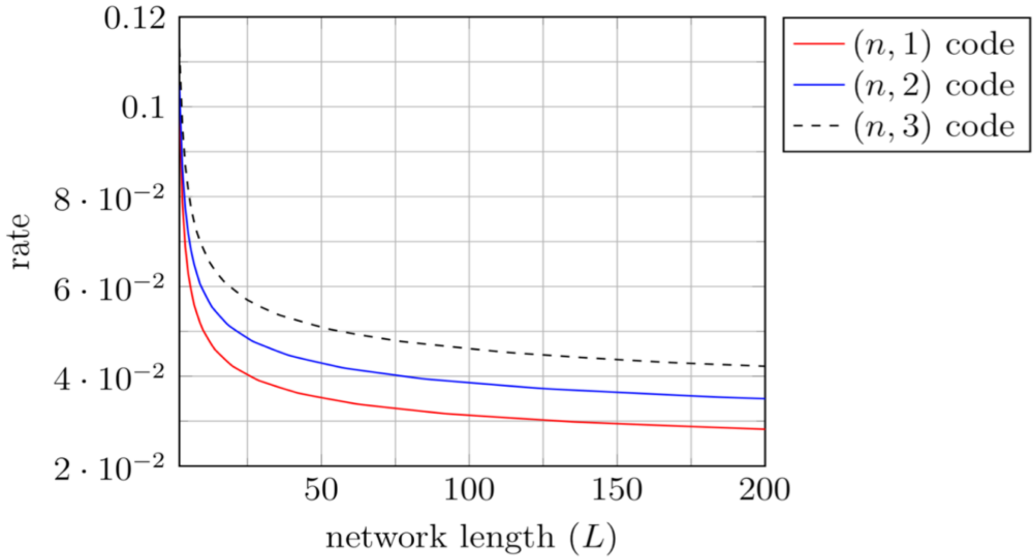

Fig. 1: Achievable rates of batched codes for line networks of binary symmetric channels, where the inner code is formed by binary (n, k) codes, where $k = 1,2,3$.

References:

[1] D. Slepian,"A class of binary signaling alphabets," *Bell Syst. Tech. J.*, vol. 35, no. 1, pp. 203–234, 1956.

[2] J. Cordaro and T. Wagner, "Optimum (n, 2) codes for small values of channel error probability (corresp.)," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 349–350, Apri 1967.

[3] P.-N. Chen, H.-Y. Lin, and S. M. Moser, "Optimal ultrasmall block- codes for binary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7346–7378, Nov. 2013.

[4] Y. Dong and S. Yang, "On optimal finite-length binary codes of four codewords for binary symmetric channels," accepted by ISITA'20. Full version online: http://arxiv.org/abs/2006.07603

[5] S. Yang and J. Wang, "Upper bound scalability on achievable rates of batched codes for line networks," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, USA, Jun. 2020.

[6] S. Yang, J. Wang, Y. Dong, and Y. Zhang, "On the capacity scalability of line networks with buffer size constraints," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jun. 2019.

[7] U. Niesen, C. Fragouli, and D. Tuninetti, "On capacity of line networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4039– 4058, Nov. 2007.

[8] S. Yang and R. W. Yeung, "Batched sparse codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5322–5346, Sep. 2014.