1. Introduction

Network coding was first introduced in Ahlswede et al.'s seminal paper [ACLY00]. Network coding gives an advantage in increasing throughput of information transmission in comparison with simple routing methods for a communication network [LYC03, HKM$^+$03]. The considered communication network was a directed graph with nodes connected with each other by multiple links. The throughput gain is achieved in network coding, since the nodes are allowed to forward a *function* of their received packets, while in routing such packets can only be forwarded to another node. Kötter and Médard provided an algebraic formulation for a *linear network coding* problem and its scalar solvability. If functions of the packets on the links of the network, then we obtain a linear network coding solution, and a *solution* is an assignment of these functions such that all destination nodes can recover all of their requested messages transmitted from a source node. Network coding was further developed with *vector network coding* by Ebrahimi [EF11], where all packets are vectors of length t. In [EW18], Etzion and Wachter-Zeh proved that vector network coding based on subspace codes outperforms linear network coding for several generalizations of the well-known combination networks [RA06]. It gives the motivation for our study in this thesis on vector network coding, especially the study of *gap* measuring the difference in *alphabet sizes* between solutions of scalar and vector network coding. The alphabet size is an important parameter determining the amount of computation performed at each node [EW18]. In Section 3.3, we show that smaller alphabet sizes can be achieved by vector network coding for *generalized combination networks* (GCN) [EW18], which allows higher number of destination nodes to be connected to the network in comparison with scalar network coding.

Outline

In **Chapter 2**, we recall coding-theory-specific notions and give an introduction to the known codes that we consider in this thesis. We first give the definition of *maximum rank distance* (MRD) code and its properties. This code was mainly used to study vector solutions for several families of the GCN in [EW18]. Then, we give the definition of Grassmannian code, Covering Grassmannian code, Multiple Grassmannian code and the notion of the maximum size of a Multiple Grammanian code. Because Grassmannian codes contain subspaces of the same dimension over a finite field $\mathbb{F}_q$, they have been recently applied in the study of network coding problems, such as [EW16, EKOÖ18, EW18, EZ19]. The remaining chapters contain our study of new gap sizes for GCN. We divided them into three main parts: Chapter 3 and 4 explain how we transfer vector network coding problems for GCN to problems of finding matrices or Grassmanian codes, Chapter 5 contains new gap sizes for 2 families of GCN with combinatorial proofs based on Lovász Local Lemma (LLL), and Chapter 6 and 7 contains new computational results of vector solutions outperforming scalar solutions for the $(\epsilon = 1, \ell = 1) - \mathcal{N}_{h=3,r,s=4}$ network. The details of each chapter are mentioned below.

In **Chapter 3** and **Chapter 4**, we represent network as a matrix channel and introduce an advantage of vector solutions in alphabet sizes in comparison with scalar solutions for network coding problems. We firstly recall the motivation of network coding, and secondly we introduce how we approach such problems by fomulating the relationship between source's messages and receiver's packets by linear equation systems. Thirdly, we explain why we choose GCN for our study, and we recall separately known theorems on an existence of a scalar or a vector solution for GCN. Finally, we formulate the *gap* to measure the difference in alphabet sizes between a vector solution and a corresponding optimal scalar solution. We list known gaps for some instances of GCN in previous studies together with our new found gaps in this study.

In **Chapter 5**, we present new gaps found by combinatorial approaches based on LLL. We begin this chapter with a simple network, namely the $(\epsilon = 1, \ell = 1) - \mathcal{N}_{h=3,r,s=4}$ network, and the gap of this network is first found in our study. During the proof of the gap, we also prove there always exists vector solutions for the network, if and only if the number r of intermediate nodes is less than or equal to a certain number. After achieving the gap for the $(\epsilon = 1, \ell = 1) - \mathcal{N}_{h=3,r,s=4}$ network, we develop the proofs further for the $(\epsilon = 1, \ell = 1) - \mathcal{N}_{h,r,s}$ network and the $(\epsilon = 1, \ell > 1) - \mathcal{N}_{2\ell,r,2\ell+1}$ network. Knowing the gap motivates us to search for vector solutions achieving such gap, which leads to computational results presented in Chapter 6.

Chapter 6 shows the core steps of 4 different computational approaches to find vector solutions outperforming the optimal scalar solutions for the $(\epsilon = 1, \ell = 1) - \mathcal{N}_{h=3,r,s=4}$ with $t = 2$ and $t = 3$. We have found vector solutions of 89 nodes and 166 nodes, while the scalar solution of such network exists if and only if $r \leq 42$ and $r \leq 146$ respectively. We then conclude the new bound on maximum size of Grassmanian codes for the network, $89 \leq \mathcal{A}_2\,(6, 4, 3; 2) \leq 126$ and $166 \leq \mathcal{A}_2\,(9, 6, 3; 2) \leq 537$. While writing this thesis, the bound of $\mathcal{A}_2\,(6, 4, 3; 2)$ has been improved in [EKOÖ18] and the result of $t = 3$ have not yet been found in any studies.

The thesis is concluded in **Chapter 7.**