Linear Information Theory and its Application to the Coded Caching Problem

by

Amirhossein Tootooni Mofrad

BASc. Computer Engineering, The University of British Columbia, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Computer Science)

The University of British Columbia
(Vancouver)

April 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Linear Information Theory and its Application to the Coded Caching Problem

submitted by **Amirhossein Tootooni Mofrad** in partial fulfillment of the requirements for the degree of **Master of Science** in **Computer Science**.

Examining Committee:

Joel Friedman, Professor, Computer Science, UBC
Supervisor

Sathish Gopalakrishnan, Associate Professor, Electrical and Computer Engineering, UBC
Additional Examiner

Abstract

We investigate the peculiar properties of information theory when all random variables involved are linear functions of a given source endowed with the structure of a $\mathbb{Z}/2\mathbb{Z}$-vector space. We describe the coded caching problem and review the progress made in particular cases of the problem. We devise a new caching scheme for a special instance of the problem and extend some approaches in the coded caching literature to the linear case such that we can use our results in linear algebra to drive new lower bounds. We show one example within coded caching, with the linearity assumption, where our results in information theory lead to stronger bounds. Our results infer certain "qualitative" aspects of caching schemes, using the information they must contain, rather than giving a complete analysis of such schemes.

Lay Summary

Studying measures of information of distributions associated with random variables is fundamental to information theory. With the assumption that the random variables follow a particular linearity constraint, we can study information theory with techniques from linear algebra. This assumption allows one to quantify measures of information between random variables that are not present in the current literature. In this work, we demonstrate the properties of this *linear* information theory.

Further, we apply our results to a practical problem in network engineering called coded caching. Broadly speaking, coded caching asks: what information should be stored in the nodes of a network such that a central server can broadcast the least amount of information? Is there a lower bound on the information content of the server broadcast? We introduce a different approach to coded caching that improves the current lower bounds under certain assumptions for a specific instance of the problem.

Preface

The entire work presented in this thesis is original, as-of-yet unpublished, independent research conducted by the author, Amir Tootooni and Dr. Joel Friedman. Formulation of linear information theory was done by Dr. Friedman, in conjunction with the author of this thesis. The application of linear information theory to the coded caching problem was done by the author of this thesis and Dr. Friedman.

Table of Contents

List of Figures

Acknowledgments

I had the great privilege of working with Joel Friedman. I am forever grateful to him for his kindness, guidance, and calmness. This work would not have been possible without him. I must thank my first academic mentor, Sathish Gopalakrishnan, for his advice and collaboration throughout my undergraduate and graduate years.

I thank my beloved friends for their support and encouragement. I thank my family for their unconditional love and sacrifice without expectation.

Lastly, I thank my dear mother, who loves me, supports me, and always believes in me.

To my mother, Nasim, and my late grandfather, Mahmood

Chapter 1

Introduction

In this thesis we develop some foundations of linear algebra to create new tools in information theory one might call *linear information theory*. We consider a source of information that is described as a certain number of bits, and a problem that involves random variables of this source. We will show that if each random variable is linear, i.e., can be expressed as a set of linear functions of these sets of bits, then certain concepts result in seemingly new information theory equalities and inequalities; our main result concerns a fundamental invariant of three linear random variables—their *discoordination*. This discoordination is, more generally, a fundamental invariant of any three subspaces of a finite-dimensional vector space over an arbitrary field.

We concretely tie these tools to give a new lower bound in a special case of an open question in an area known as *coded caching*. Our form of the problem is a mild restriction of the original. Our bounds in coded caching will use the aforementioned discoordination. For the rest of this chapter we summarise our main ideas and results.

1.1 The Linearity Assumption

In information theory, a random variable is defined as a map $Y\colon S \to \mathcal{Y}$ where S is a finite set with a probability measure $P\colon S \to \mathbb{R}$ and $\mathcal{Y}$ is a finite set. If we assume that both S and $\mathcal{Y}$ are finite-dimensional $\mathbb{F}$-vector spaces and Y is a linear map,

then Y can be identified with a unique linear subspace, U, in the dual space of S. We refer to U as the *linear random variable* corresponding to Y. In this case, for $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ and a uniformly distributed Y we have

$$H_2(U) := H_2(Y) = \dim(U),$$

where $H_2(Y)$ denotes the base 2 entropy of the random variable Y and we define $H_2(U)$ to be equal to $H_2(Y)$. Furthermore, for random variables Y_1 and Y_2 with the same assumptions as above and corresponding linear random variables $U_1, U_2 \subset S^*$, we have

$$H_2(U_1, U_2) := H_2(Y_1, Y_2) = \dim\big(\mathrm{Span}(U_1, U_2)\big),$$

where we define $H_2(U_1, U_2)$ to equal $H_2(Y_1, Y_2)$. This linearity assumption is the bridge between information theory and the linear algebra developed in this work. In essence, linear information theory is when quantities such as entropy and mutual information can be represented in relation to the dimension of linear subspaces.

1.2 Main Results in Linear Algebra

We define $I(Y_1;Y_2)$ and $I(Y_1;Y_2;Y_3)$ as the usual two-way and three-way *mutual information* of the random variables Y_1, Y_2, Y_3; namely

$$I(Y_1;Y_2) = H(Y_1) + H(Y_2) - H(Y_1, Y_2),$$

$$I(Y_1;Y_2;Y_3) = I(Y_1;Y_2) + I(Y_1;Y_3) - I(Y_1;Y_2,Y_3).$$

If Y_1, Y_2, Y_3 follow the linearity assumption from Subsection 1.1, then U_1, U_2, U_3 are their respective corresponding linear random variables which are linear subspaces of an ambient finite-dimensional vector space, $\mathcal{U}$, and

$$I(Y_1;Y_2) = I(U_1;U_2), \quad I(Y_1;Y_2;Y_3) = I(U_1;U_2;U_3).$$

The linear algebra we develop generalizes what is often called the "dimension formula," [2, 6] which states,

$$\dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 + U_2),$$

here $U_1 + U_2$ denotes the sum (equal to the span) of U_1 and U_2 in the ambient vector space $\mathcal{U}$. Of course, the vector spaces $U_1 \cap U_2$ and $U_1 + U_2$ are not intrinsic to the isomorphism class U_1 and U_2, but depend on their relation to the ambient vector space. The dimension formula shows $I(U_1;U_2)$ equals $\dim(U_1 \cap U_2)$.

In contrast to the dimension formula, it is well known that $I(U_1;U_2;U_3)$ does not generally equal $\dim(U_1 \cap U_2 \cap U_3)$, see Exercise 9, Section 3.3 on page 51 of Jänich [6]. Equality does hold if the U_1, U_2, U_3 are *coordinated* in the sense that they have a *coordinating basis*, meaning a basis, X, of $\mathcal{U}$, such that for $i = 1,2,3$ the vectors $X \cap U_i = \{x \in X \,|\, x \in U_i\}$ span U_i.

A simple example where $I(U_1;U_2;U_3)$ does not equal $\dim(U_1 \cap U_2 \cap U_3)$ is where $\mathcal{U} = \mathbb{F}^2$, for an arbitrary field $\mathbb{F}$, and

$$U_1 = \mathrm{Span}(e_1),\ U_2 = \mathrm{Span}(e_2),\ U_3 = \mathrm{Span}(e_1 + e_2), \tag{1.2.1}$$

here e_1, e_2 are the standard basis vectors, in which case $\dim(U_1 \cap U_2 \cap U_3) = 0$ but $I(U_1;U_2;U_3) = -1$.

Fundamentally, we show (1.2.1) is essentially *the only* example where $I(U_1;U_2;U_3)$ does not equal $\dim(U_1 \cap U_2 \cap U_3)$.

More precisely, if $U_1, U_2, U_3 \subset \mathcal{U}$ are three subspaces of a finite dimensional $\mathbb{F}$-vector space, $\mathcal{U}$, then we may decompose $\mathcal{U}$ as a direct sum of subspaces $\mathcal{U}_1$ and $\mathcal{U}_2$, such that the restriction of U_1, U_2, U_3 to $\mathcal{U}_1$ are coordinated there, and there is an isomorphism $\iota : \mathcal{U}_2 \to \mathbb{F}^2 \otimes \mathbb{F}^m$ for some $m \geq 0$, under which ι applied to the restriction of the U_1, U_2, U_3 to $\mathcal{U}_2$ is

$$\begin{aligned} \iota\left(U_1\big|_{\mathcal{U}_2}\right) &= \mathrm{Span}(e_1) \otimes \mathbb{F}^m, \\ \iota\left(U_2\big|_{\mathcal{U}_2}\right) &= \mathrm{Span}(e_2) \otimes \mathbb{F}^m, \\ \iota\left(U_3\big|_{\mathcal{U}_2}\right) &= \mathrm{Span}(e_1 + e_2) \otimes \mathbb{F}^m. \end{aligned} \tag{1.2.2}$$

The integer m is uniquely determined and what we define as the *discoordination* of U_1, U_2, U_3, denoted $\mathrm{DisCoord}^{\mathcal{U}}(U_1, U_2, U_3)$; one can give a number of equivalent definitions of this integer m.

In addition to the discoordination of three subspaces of an ambient vector

space, we give some theorems on families of subspaces that are coordinated, and study how $\mathrm{DisCoord}^{\mathcal{U}}(U_1,U_2,U_3)$ changes when replacing $\mathcal{U}$ with a quotient space, $\mathcal{U}/W$, of $\mathcal{U}$, whereupon we consider the images of the U_i in this quotient space, i.e., we replace U_i with the equivalence classes of U_i+W (or $[U_i]_W$) and consider $\mathrm{DisCoord}^{\mathcal{U}/W}([U_1]_W,[U_2]_W,[U_3]_W)$.

1.3 Expressiveness of Linear Information Theory

Linear information theory seems more expressive than classical information theory, in that for linear random variables $U_1,U_2,U_3 \subset \mathcal{U}$, we have

$$I(U_1;U_2;U_3) = \dim(U_1 \cap U_2 \cap U_3) - \mathrm{DisCoord}^{\mathcal{U}}(U_1,U_2,U_3).$$

With the linearity assumption we can express the "non-negative" and "non-positive" parts of the mutual information among three subspaces.

Furthermore, we are able to quantify $\dim^{\mathcal{U}/U_3}([U_1 \cap U_2]_{U_3})$, which is the dimension of the image of $U_1 \cap U_2$ in $\mathcal{U}/U_3$. This is not always equal to $I(U_1;U_2\,|\,U_3)$, as in our context we have

$$I(U_1;U_2|U_3) = \dim^{\mathcal{U}/U_3}([U_1]_{U_3} \cap [U_2]_{U_3}),$$

which is the dimension of the intersection of the images of U_1 and U_2 in $\mathcal{U}/U_3$. As far is we know, it is not possible to express the following quantities in classical information theory

$$H(U_1 \cap U_2 \cap U_3) := \dim(U_1 \cap U_2 \cap U_3), \quad H^{\mathcal{U}/U_3}(U_1 \cap U_2) := \dim^{\mathcal{U}/U_3}([U_1 \cap U_2]_{U_3}).$$

1.4 Overview of Coded Caching

The second part of this thesis shows how one can apply the discoordination of three subspaces and improve upon the current results in the "easiest" open problem in a class of problems in information theory known collectively as *coded caching*.

What makes this prototypical problem in this field intriguing—its application to network engineering aside—is that with minor simplifications it becomes

a mathematical puzzle, and there have been several different mathematical approaches to tackle this problem.

The easiest open problem in which we were able to apply our results from linear algebra can be informally stated as such: Consider three students, S_1, S_2, S_3 in a course, each of whom must write an essay for a course, each essay requiring the specialized knowledge that can be found in exactly one book in the university's electronic library. Each book is written in a binary alphabet, consisting of F bits of information. At Tuesday noon, the instructor has not decided on the essay topics, but has decided on the three books, B_1, B_2, B_3, which the topics will be based on; the instructor posts the list of these three books at noon, and intends to post the list of essays the next day at noon. Tuesday night, is a time of "low usage" on the network, meaning each student can download all three books and perform any calculations they wish, but by 9 am on Wednesday they can only devote $MF \leq 3F$ bits of storage on their laptops, for some real number M. The instructor posts the essay topics at Wednesday noon, which is a time of "high usage" of the network; at this point, each student immediately chooses an essay topic (or the topics could be assigned randomly by the instructor, etc.) and broadcasts their choice of essay (in actuality the book corresponding to the essay topic) to the library; hence the library receives a function $\sigma\colon \{S_1, S_2, S_3\} \to \{B_1, B_2, B_3\}$ with $\sigma(S_i)$ denoting the book required by student S_i. Each student wants to begin working on their essay on Wednesday afternoon, however, due to high network usage, the library can only broadcast $RF \leq 3F$ bits of information, for some real number R. The question is for which values of R do there exist Z_1, Z_2, Z_3, each $\{0,1\}^{3F} \to \{0,1\}^{MF}$, with Z_i representing what student S_i stores in their cache, such that for any σ there exist a function $g_\sigma\colon \{0,1\}^{3F} \to \{0,1\}^{RF}$, such that for each Z_i and g_σ, student S_i can determine all F bits in the book $\sigma(S_i)$.

1.5 Thesis Organization

This thesis is organized as follows:

In Chapter 2, we review the notation and linear algebra concepts used in the work. We formalize the linearity assumption relating coded caching to our results in linear algebra. We give a proof of the dimension formula while highlighting its

relation to our results in linear algebra.

In Chapter 3, we formally define "discoordination" and state our main results about coordination and discoordination. This chapter establishes properties of (dis)coordination and proves the main results.

In Chapter 4, we formally state the coded caching problem and review the relevant literature. We build upon the work done by Tian (in [9]) regarding a small instance of the coded caching problem and derive a new bound using his approach. We present a new caching scheme in Section 4.6 which achieves a new memory-rate pair not present in the literature. Lastly, we use our results from Chapter 3 to get a stronger bound for a specific instance of the coded caching problem.

In Chapter 5, we give an overview and suggest directions for future research. Lastly, the author of this thesis attempts to give the context and story behind this work in Section 5.3.

Chapter 2

Linear Algebra and Information Theory Preliminaries

In this chapter, we will give some basic linear algebra notation and conventions used to define "coordination" and "discoordination" of subspaces of a fixed ambient vector space. We refer to [2, 6] for basic notions in linear algebra, quotient vector spaces, etc. We will briefly review these as needed. We also review some relevant concepts from information theory and define linear random variables.

2.1 Set Theoretic Notation

We use $\mathbb{Z}$ and $\mathbb{R}$ to respectively denote the integers and the real numbers. We use $\mathbb{N}$ to denote the natural numbers $\{1,2,\ldots\}$, and for $n \in \mathbb{N}$ we use $[n]$ to denote $\{1,\ldots,n\}$. If A,B are sets we use $A \setminus B$ to denote the set difference of A and B, meaning $A \setminus B = \{a \in A \mid a \notin B\}$.

2.2 Inequality Summation Principle

The following trivial proposition is used in a number of arguments in this work and is surprisingly useful (it is the idea behind *complementary slackness* in linear programming), we refer to it as the "Inequality Summation Principle."

Proposition 2.2.1. *For $m \in \mathbb{N}$, consider m inequalities*

$$s_1 \leq t_1, \ \cdots, \ s_m \leq t_m \tag{2.2.1}$$

that hold for real numbers $s_1, \ldots, s_m$ and $t_1, \ldots, t_m$, then

$$s_1 + \cdots + s_m \leq t_1 + \cdots + t_m, \tag{2.2.2}$$

and equality holds in (2.2.2) *iff equality holds in all the inequalities in* (2.2.1)*.*

Proof. The proof is immediate: if equality holds in all inequalities of (2.2.1), then it holds in (2.2.2); otherwise, at least one inequality in (2.2.1) is strict, whereupon (2.2.2) must be strict. □

2.3 Algebra Notation

2.3.1 Direct Sum and $\oplus$

In mathematics, $\oplus$ usually denotes the direct sum of vector spaces. However, in the coded caching literature, $\oplus$ is usually used for the addition of vectors in a vector space over the field $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$. In order to avoid confusion, we will use $\underline{\oplus}$ for the direct sum of vector spaces and keep the coded caching convention unchanged.

Let U_1, U_2 be finite dimensional subspaces of a vector space $\mathcal{U}$ over $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ with $\dim(U_1) = \dim(U_2)$, and let $\mu \colon U_1 \to U_2$ be an isomorphism. Then we use $U_1 \oplus_\mu U_2$ to denote the subspace of $\mathcal{U}$ consisting of all vectors $u_1 + \mu(u_1)$ with $u_1 \in U_1$. Often μ will be understood (or unimportant), in which case we write $U_1 \oplus U_2$. Hence $U_1 \oplus U_2$ always connotes that there is an understood isomorphism $U_1 \to U_2$. In our case these isomorphisms, such as those in Lemma 4.5.2 are built by choosing an ordered basis $a_1, \ldots, a_k$ for U_1 and another $b_1, \ldots, b_k$ for U_2, and picking $\mu \colon U_1 \to U_2$ as the unique linear map $\mu(a_i) = b_i$ for $i \in [m]$.

Similarly, if U_1, U_2, U_3 are finite dimensional subspaces of a vector space $\mathcal{U}$ over $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ and $\mu_i \colon U_1 \to U_i$ are isomorphisms for $i = 2, 3$, then

$$U_1 \oplus_{\mu_2} U_2 \oplus_{\mu_3} U_3$$

denotes the subspace of $\mathcal{U}$ given by all vectors $u_1 + \mu_2(u_1) + \mu_3(u_1)$ with $u_1 \in U_1$. We will use this notation in Lemma 4.5.2 and the discussion that follows it in Chapter 4.

2.3.2 $\mathbb{F}$-Universes

Definition 2.3.1. Let $\mathbb{F}$ be an arbitrary field. By an $\mathbb{F}$-*universe*, $\mathcal{U}$, we mean a finite-dimensional $\mathbb{F}$-vector space. By the term "$\mathbb{F}$-universe," without mention of $\mathbb{F}$, we understand that $\mathbb{F}$ is an arbitrary field.

In this work the field $\mathbb{F}$ and $\mathbb{F}$-universe, $\mathcal{U}$, are generally fixed or, at least, understood in context. Hence notions such as "a subspace of $\mathcal{U}$" and its "dimension" can be used unambiguously. When we are working with more than one ambient vector space we will use $\dim^{\mathcal{U}}(U)$ to denote the dimension of the subspace U in $\mathcal{U}$.

At times we use results that hold when the ambient vector space $\mathcal{U}$ can be infinite dimensional; in this case we use the term "$\mathbb{F}$-vector space," and, similarly, we understand $\mathbb{F}$ to be an arbitrary field unless explicitly mentioned otherwise. However, in this work we mostly limit ourselves to ambient vector spaces, $\mathcal{U}$, that are finite dimensional.

2.3.3 Sum and Span

If A, B are subsets of an $\mathbb{F}$-vector space, $\mathcal{U}$, the *sum of A and B* refers to the set

$$A + B = \{a + b \mid a \in A, b \in B\}; \tag{2.3.1}$$

if $U_1, U_2 \subset \mathcal{U}$ are subspaces, then so is $U_1 + U_2$; we similarly define $S_1 + \cdots + S_m$ for any subsets $S_1, \ldots, S_m$ of $\mathcal{U}$. If $S_1, \ldots, S_m$ are subsets of $\mathcal{U}$, we use

$$\mathrm{Span}(S_1, \ldots, S_m)$$

to denote the span of $S_1, \ldots, S_m$; if $S_1, \ldots, S_m$ are subspaces, then this span equals $S_1 + \cdots + S_m$.

Definition 2.3.2. For an $\mathbb{F}$-vector space, $\mathcal{U}$, let $\mathrm{Ind}(\mathcal{U})$ denote the set of all sets of linearly independent vectors in $\mathcal{U}$, meaning for any $X \in \mathrm{Ind}(\mathcal{U})$, X is a set of

linearly independent vectors in $\mathcal{U}$ that spans a subspace of dimension $|X|$.

2.4 Quotient Space and Relative Basis Conventions

First we recall the usual notion of a *quotient space* of vector spaces; see [2, 6] for details. For an $\mathbb{F}$-universe, $\mathcal{U}$, let $u_1, u_2 \in \mathcal{U}$ be vectors, and $W \subset \mathcal{U}$ be a subspace. By a *W-coset of $\mathcal{U}$* we mean any set of the form $u + W$ where $u \in \mathcal{U}$ and $+$ as in (2.3.1); it is convenient to denote $u + W$ by $[u]_W$, and we use $\mathcal{U}/W$ to denote the set of all W-cosets. We have that $u_1 + W = u_2 + W$ iff $u_1 - u_2 \in W$, so one can view $\mathcal{U}/W$ as the set of equivalence classes under the equivalence $u_1 \sim u_2$ iff $u_1 - u_2 \in W$. It is easy to check that the vector space structure on W and $\mathcal{U}$ gives rise to one on $\mathcal{U}/W$, and that

$$\dim(\mathcal{U}/W) = \dim(\mathcal{U}) - \dim(W).$$

If $Y \subset \mathcal{U}$ is any subset of $\mathcal{U}$, we use the notation $[Y]_W$ to denote the set of W-cosets $Y + W$, viewed as a subset of $\mathcal{U}/W$; we call $[Y]_W$ the *image of Y in $\mathcal{U}/W$* (used in Section 1.2); hence if $y \in Y$, then $[\{y\}]_W$ is the one element set $[y]_W \in [Y]_W$.

If $\mathcal{U}$ is an $\mathbb{F}$-universe and $W \subset \mathcal{U}$ is some subspace in $\mathcal{U}$, part of our methods examines what happens to certain subspaces of $\mathcal{U}$ when we consider their image in $\mathcal{U}/W$. Further, if $U \subset \mathcal{U}$ is another subspace, then $[U]_W$ is a subspace of $\mathcal{U}/W$, but (we can easily check that) $[U]_W$ is isomorphic to the image of U in $\mathcal{U}/(U \cap W)$. Hence

$$\dim^{\mathcal{U}/W}([U]_W) = \dim^{\mathcal{U}}(U) - \dim^{\mathcal{U}}(U \cap W)$$

which equals $\dim(U) - \dim(W)$ only when $W \subset U$. At times we write U/W to denote the image of U in $\mathcal{U}/W$. In some instances, for ease of notation we may write $\dim(U/W)$, $\dim^{\mathcal{U}/W}(U)$, or $\dim^{\mathcal{U}/W}([U])$ instead of $\dim^{\mathcal{U}/W}([U]_W)$.

Definition 2.4.1. Let $W \subset U$ be subspaces of a vector space, $\mathcal{U}$, such that $\dim(W) = m$ and $\dim(U) = n$. We say that a subset, $Y = \{y_1, \ldots, y_{n-m}\}$, of U is a *basis of U relative to W* if the image of Y in $\mathcal{U}/W$, i.e., $[Y]_W = \{y_1 + W, \ldots, y_{n-m} + W\}$, is a basis of the image of U in $\mathcal{U}/W$ (i.e., $[U]_W$).

By Definition 2.4.1 we have that, if $X = \{x_1, \ldots, x_m\}$ is any basis for W, then $Y = \{y_1, \ldots, y_{n-m}\}$ is a basis of U relative to W iff $X \cup Y$ is a basis for U. While

we may think of Y as what we add to X to complete the basis, the above definition shows that our choice of Y depends only on $W = \mathrm{Span}(X)$ and not X itself.

2.5 Independent Subspaces, Decompositions, and Factorization

The notion of the linear independence of subspaces of a vector space is not a standard one, even though it likely occurs implicitly in the literature.

Consider subspaces $U_1, \ldots, U_m$ of an $\mathbb{F}$-universe, for each $i \in [m]$, let X_i be a basis for U_i and $u_i \in U_i$. Then each vector in $U_1 + \cdots + U_m$ can be written as $u_1 + \cdots + u_m$, and hence lies in the span of $X_1 \cup \cdots \cup X_m$. Therefore,

$$\dim(U_1 + \cdots + U_m) \leq |X_1 \cup \cdots \cup X_m| \leq |X_1| + \cdots + |X_m|,$$

consequently

$$\dim(U_1 + \cdots + U_m) \leq \dim(U_1) + \cdots + \dim(U_m); \tag{2.5.1}$$

furthermore, strict inequality holds in one of two cases:

1. the $X_1, \ldots, X_m$ are not distinct; or
2. some proper subset of $X_1 \cup \cdots \cup X_m$ also spans $U_1 + \cdots + U_m$.

Both cases imply that for some $i \in [m]$, and some $x \in X_i$, x can be expressed as a linear combination of the vectors in $X_i \setminus \{x\}$ and the remaining X_j such that $j \neq i$. Since the vectors in each of the bases are linearly independent, this expression leads to an equation

$$u_1 + \cdots + u_m = 0 \quad \text{and } u_i \neq 0,$$

where $u_i \in U_i$ for all $i \in [m]$. Conversely, if equality holds in (2.5.1), then $X_1, \ldots, X_m$ are necessarily distinct and their union is a linearly independent set that spans $U_1 + \cdots + U_m$; hence this union comprises a basis for $U_1 + \cdots + U_m$. There are numerous equivalent ways in which (2.5.1) holds with equality, they are minor variants of conditions given in the following definition.

Definition 2.5.1. Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. We say that $U_1, \ldots, U_m$ are *linearly independent* if any of the following conditions hold.

1. For all $u_1, \ldots, u_m$ with $u_i \in U_i$, if $u_1 + \cdots + u_m$ is 0, then $u_i = 0$ for all $i \in [m]$.
2. Any $u \in U_1 + \cdots + U_m$ has a unique representation as a sum $u = u_1 + \cdots + u_m$ with $u_i \in U_i$.
3. For any bases $X_1, \ldots, X_m$ of $U_1, \ldots, U_m$ respectively, the $X_1, \ldots, X_m$ are mutually disjoint and $X_1 \cup \cdots \cup X_m$ is a basis for $U_1 + \cdots + U_m \subset \mathcal{U}$.
4. There exist bases $X_1, \ldots, X_m$ of $U_1, \ldots, U_m$ respectively, such that the $X_1, \ldots, X_m$ are mutually disjoint and $X_1 \cup \cdots \cup X_m$ is a basis for $U_1 + \cdots + U_m \subset \mathcal{U}$.
5. $$\dim(U_1) + \cdots + \dim(U_m) = \dim(U_1 + \cdots + U_m). \tag{2.5.2}$$

We note that condition (5) makes use of the fact that $\mathcal{U}$ is finite dimensional, whereas (1)–(4) above are equivalent when $\mathcal{U}$ is any $\mathbb{F}$-vector space such that any subspace of $\mathcal{U}$ has a basis[1].

Example 2.5.2. If $u_1, \ldots, u_m$ are vectors in some vector space, the vectors are linearly independent iff all these vectors are nonzero and $\mathrm{Span}(u_1), \ldots, \mathrm{Span}(u_m)$ are linearly independent subspaces. Thus the classical notion of linearly independent vectors can be described in terms of the linear independence of one-dimensional subspaces.

Example 2.5.3. If $B_1, \ldots, B_m$ is any partition of a set of linearly independent vectors in a any vector space, then their spans are linearly independent subspaces.

Just as in Definition 2.5.1 we can easily check that the three conditions in the following definition are equivalent.

Definition 2.5.4. By a *decomposition* of a subspace $U \subset \mathcal{U}$ of an $\mathbb{F}$-universe, $\mathcal{U}$, we mean subspaces $U_1, \ldots, U_m \subset U$ such that any of these equivalent conditions hold.

[1] Such a condition is typically assumed in linear algebra, although depending on the vector spaces this may require a set theoretic assumption such as transfinite induction.

1. Each $u \in U$ can be written uniquely as $u_1 + \cdots + u_m$ where $u_i \in U_i$ for all $i \in [m]$.

2. The $U_1, \ldots, U_m$ are linearly independent and their span is all of U.

3. The map $U_1 \underline{\oplus} \cdots \underline{\oplus} U_m \to U$ defined by $(u_1, \ldots, u_m) \mapsto u_1 + \cdots + u_m$ is an isomorphism.

For some subspace $A \subset U$, the first condition in Definition 2.5.1 implies that if $U_1, \ldots, U_m$ are any linearly independent subspaces, then so are $A \cap U_1, \ldots, A \cap U_m$.

Definition 2.5.5. Let $U_1, \ldots, U_m$ be a decomposition of a subspace U of some $\mathbb{F}$-universe. We say that a subspace $A \subset U$ *factors through* this decomposition if any of these equivalent conditions hold.

1. $A \cap U_1, \ldots, A \cap U_m$ is a decomposition of A.

2. Any vector in A can be written as a unique sum of vectors in $A \cap U_1, \ldots, A \cap U_m$.

3. The span of $A \cap U_1, \ldots, A \cap U_m$ is all of A.

4. $$\sum_{i=1}^{m} \dim(A \cap U_i) = \dim(A).$$

The following proposition likely occurs in a number of places in the literature.

Proposition 2.5.6. *If $A, B \subset U$ factor through a decomposition $U_1, \ldots, U_m$ of a subspace, U, of some universe then $A + B$, $A \cap B$ also factor through this decomposition.*

Proof. For any $i \in [m]$, by the dimension formula and since $A \cap U_i, B \cap U_i$ are both subspaces of $(A + B) \cap U_i$ we have

$$\begin{aligned} \dim(A \cap U_i) + \dim(B \cap U_i) &= \dim(A \cap B \cap U_i) + \dim\big((A \cap U_i) + (B \cap U_i)\big) \\ &\leq \dim(A \cap B \cap U_i) + \dim\big((A + B) \cap U_i)\big). \end{aligned}$$

Summing this inequality over all over all i we have

$$\dim(A\cap B)+\dim(A+B)\leq\sum_{i=1}^{m}\dim(A\cap B\cap U_i)+\sum_{i=1}^{m}\dim\big((A+B)\cap U_i)\big). \quad (2.5.3)$$

Given that the $(A\cap B)\cap U_i$ are linearly independent subspaces of $A\cap B$, and the $(A+B)\cap U_i$ are linearly independent subspaces of $A+B$, by (2.5.2) we have

$$\sum_{i=1}^{m}\dim\big((A\cap B)\cap U_i\big)=\dim\big(\sum_{i=1}^{m}(A\cap B)\cap U_i\big)\leq A\cap B,$$
$$\sum_{i=1}^{m}\dim\big((A+B)\cap U_i\big)=\dim\big(\sum_{i=1}^{m}(A+B)\cap U_i\big)\leq A+B.$$

Summing the above two inequalities shows that (2.5.3) holds with equality. □

From the above proposition we get the following useful consequence.

Theorem 2.5.7. *Let $U_1,\ldots,U_m$ be a decomposition of a subspace U of some universe. Say that each of the subspaces $A_1,\ldots,A_s\subset U$ factors through this decomposition. Then any subspace that can be written as an expression involving $+$ and $\cap$ and the $A_1,\ldots,A_s$ (and parenthesis) factors through this decomposition.*

One can prove the above theorem by induction on the "size" of the expression where by size we mean the number of the A_i, $+$, and $\cap$ present in the expression.

Similarly, note that if A,B factor through such a decomposition, then

$$\dim(A/B)=\dim\big(A/(A\cap B)\big)=\dim(A)-\dim(A\cap B)$$

$$=\sum_{i=1}^{m}\dim(A\cap U_i)-\sum_{i=1}^{m}\dim(A\cap B\cap U_i)=\sum_{i=1}^{m}\dim\big((A\cap U_i)/(A\cap B\cap U_i)\big).$$

Hence $\dim(A/B)$ can be computed by restricting both A and $A\cap B$ to each U_i, and computing the dimension of the quotient space $(A\cap U_i)/(A\cap B\cap U_i)$ there.

Note that the definitions, proposition, and theorem stated in this section hold when U is taken to be the whole $\mathbb{F}$-universe, $\mathcal{U}$. Meaning we can consider a decomposition $\mathcal{U}_1,\ldots,\mathcal{U}_m$ of $\mathcal{U}$ and restate all the above.

2.6 Basis Exchange and Independent Subspaces

We use numerous variants of the basis exchange and basis extension principles (see [6]). The following remark details what we need in subsequent chapters.

Remark 2.6.1. *Let U be a subspace of any $\mathbb{F}$-universe. Then*

1. *if $U = \mathrm{Span}(S)$ for some subset, S, of $\mathcal{U}$, then some subset $S' \subset S$ is a basis for U;*
2. *if X is a basis of U, $X_0 \subset X$ a subset, and Y any set of linearly independent vectors with $\mathrm{Span}(X_0)$ and $\mathrm{Span}(Y)$ linearly independent, then there exists a basis for U consisting of $X_0 \cup Y$ plus a subset of vectors from $X \setminus X_0$;*
3. *if X is a basis for U and Y a subset of linearly independent vectors in U, then there is a basis of U of the form $Y \cup X_0$ with $X_0 \subset X$ (this is the standard basis exchange principle).*

2.7 The Dimension Formula and Its Proof

Let us recall the *dimension theorem* and its proof; for more detail, see Theorem 3 in Section 3.2 (page 49) of Jänich [6] (called there the "Dimension formula for subspaces"). This illustrates the idea behind our main technique to show that certain subspaces of a universe are *coordinated* (we formally define this notion in Chapter 3).

The dimension formula states that if $U_1, U_2 \subset \mathcal{U}$ are subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, then

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim(U_1) + \dim(U_2). \tag{2.7.1}$$

The proof sketch is as follows:

1. let B_0 be a basis for $U_1 \cap U_2$;
2. extend B_0 (in an arbitrary fashion) to a basis $B_0 \cup B_1$ of U_1;
3. extend B_0 (in an arbitrary fashion) to a basis $B_0 \cup B_2$ of U_2;
4. verify that B_0, B_1, B_2 are disjoint and $B = B_0 \cup B_1 \cup B_2$ is a linearly independent set that spans $U_1 + U_2$;

5. conclude $\dim(U_1 \cap U_2) = |B_0|, \dim(U_1 + U_2) = |B_0| + |B_1| + |B_2|$ and that $\dim(U_i) = |B_0| + |B_i|$ for $i = 1, 2$

Our main theorem about coordination generalizes the above verification, let us summarize how this is done. B_1 and B_2 are extension pieces of B_0 and hence disjoint from B_0. If B_1 and B_2 intersect, say at a vector, v, then $v \in (U_1 \cap U_2) = \text{Span}(B_0)$. Therefore, v cannot be an extension piece from B_0 to $B_0 \cup B_1$ (or $B_0 \cup B_2$). Finally, if $B = B_0 \cup B_1 \cup B_2$ are not disjoint or not linearly independent, then there is a nontrivial solution to the equation

$$v_0 + v_1 + v_2 = 0$$

where each $v_i \in \text{Span}(B_i)$ and "nontrivial" means that at least one of v_0, v_1, v_2 is nonzero, hence at least two of v_0, v_1, v_2 are nonzero. This leads to the following contradiction: assuming such a nontrivial solution, at least one of v_1, v_2 is nonzero, say $v_1 \neq 0$, then we have $v_2 \in U_2$ and $v_0 \in (U_1 \cap U_2) \subset U_2$, so both $v_0, v_2 \in U_2$ and as a result

$$v_1 = \ -v_0 - v_2 \ \in U_2;$$

but since $v_1 \in U_1$ this implies

$$v_1 \in U_1 \cap U_2$$

which is impossible since v_1 is nonzero and a linear combination of vectors in B_1, v_1 cannot lie in the span of B_0 given that B_1 is an extension piece of B_0.

We took an equation $v_0 + v_1 + v_2 = 0$, wrote it as $v_1 = -v_0 - v_2$, and obtained the result $v_1 \in U_1 \cap U_2$ which was additional information about v_1 that lead to a contradiction. We will use such arguments in Section 3.3. Leveraging the generalization of such arguments is the main motivation behind our definition of the *greedy algorithm* (Subsection 3.2.9) and *quasi-increasing sequences of subspaces* (Subsection 3.3.1).

2.7.1 The Dimension Formula in Infinite Dimensions

The proof of the dimension formula in Section 2.7 has an infinite dimensional version which is also valid for $U_1, U_2 \subset \mathcal{U}$ where $\mathcal{U}$ is an $\mathbb{F}$-vector space, and one

or both of U_1, U_2 are infinite dimensional, namely

$$0 \to U_1 \cap U_2 \to U_1 \underline{\oplus} U_2 \to U_1 + U_2 \to 0$$

is an *exact sequence*, meaning that the kernel of any arrow equals the image of the preceding arrow [1].

It is therefore likely that some of our discussion regarding subspaces of an ambient $\mathbb{F}$-universe hold in greater generality. However, the above exact sequence suggests to us that it is simpler to work out our results in the finite dimensional case, and then see which results could have a useful infinite dimensional analog.

2.8 Linear Random Variables and a Review of Information Theory

Throughout this section, $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ is the finite field of two elements. Here we review the usual notion of entropy and explain what we mean by a *linear random variable* of an $\mathbb{F}$-universe.

In classical information theory (see [4] for more detail), a *source*, S, is a finite set with a probability measure $P\colon S \to \mathbb{R}$ whose values are positive and sum to one $\big($we do not allow $P(s) = 0$ for any $s \in S\big)$. A random variable is defined as a map $Y\colon S \to \mathcal{Y}$ where $\mathcal{Y}$ is a finite set. For each $y \in \mathcal{Y}$, we define

$$p_y = \sum_{s \in Y^{-1}(y)} P(s),$$

and we define its (base 2) entropy to be

$$H(Y) = H_2(Y) = \sum_{y \in \mathcal{Y}} p_y \log_2(1/p_y),$$

where $p_y \log_2(1/p_y)$ is taken to be 0 if $p_y = 0$.

If Y is *uniformly distributed* in the sense that p_y is independent of y, it follows that $p_y = 1/|\mathcal{Y}|$, and

$$H(Y) = \log_2(|\mathcal{Y}|). \tag{2.8.1}$$

Each random variable $Y\colon S \to \mathcal{Y}$ induces a partition of S, namely

$$S = \bigcup_{y\in\mathcal{Y}} Y^{-1}(y).$$

Another random variable $Y'\colon S \to \mathcal{Y}'$ is *equivalent* to Y if it gives the same partition of S; this holds iff there is an isomorphism $\mu\colon \mathrm{Image}(Y) \to \mathrm{Image}(Y')$ such that $Y' = \mu \circ Y$. Meaning, if $Y_1,\dots,Y_m$ are equivalent, respectively, to $Y_1',\dots,Y_m'$, then any expression involving the joint entropy, mutual information, conditional entropy, etc., involving the $Y_1,\dots,Y_m$ equals that when each Y_i is replaced with Y_i'.

If $Y\colon S \to \mathcal{Y}$ is any random variable, then Y is equivalent to the random variable where we discard any $y \in \mathcal{Y}$ with $p_y = 0$; since we assume that each $s \in S$ has positive probability, this is the same as discarding all elements of $\mathcal{Y}$ that are not in the image of Y. This amounts to replacing Y with the map it induces from S to $\mathrm{Image}(Y)$, which is a surjective map; we call this new random the *surjective version of* Y. If Y_1, Y_2 are surjective random variables, $Y_i\colon S \to \mathcal{Y}_i$, then Y_1 is equivalent to Y_2 iff there exists a bijection $\mu\colon \mathcal{Y}_1 \to \mathcal{Y}_2$ with $Y_2 = \mu \circ Y_1$.

Definition 2.8.1. Let $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, and let S be an $\mathbb{F}$-universe. We view S as a probability space with the uniform distribution, i.e., each element occurs with probability $1/|S| = 1/2^n$ where $n = \dim(S)$. By a *classical linear random variable* we mean a linear map $Y\colon S \to \mathcal{Y}$ where $\mathcal{Y}$ is an $\mathbb{F}$-universe.

Proposition 2.8.2. *For any classical linear random variable $Y\colon S \to \mathcal{Y}$, there exists an equivalent random variable which is a quotient map $\tilde{Y}\colon S \to S/A$ where $A = \ker(Y) = \ker(\tilde{Y})$. Furthermore,*

$$H(Y) = H(\tilde{Y}) = \log_2(|S/A|) = \dim(S/A) = \dim(S) - \dim(A). \tag{2.8.2}$$

Proof. Any linear map $Y\colon S \to \mathcal{Y}$ factors uniquely as

$$S \xrightarrow{f} S/\ker(Y) \xrightarrow{g} \mathcal{Y},$$

with f surjective and g injective. Y is equivalent to its surjective form, and if Y is surjective, then g is also surjective; in this case $g\colon S/A \to \mathcal{Y}$ is a bijection, and hence g gives an equivalence of the surjective form of Y and the map $\tilde{Y}\colon S \to S/A$

where $\ker(Y) = A$.

Since $\tilde{Y}$ is surjective and linear, it is uniform, using (2.8.1) we have

$$H(Y) = H(\tilde{Y}) = \log_2(|S/A|),$$

(2.8.2) follows. □

Proposition 2.8.3. *If $A_1, A_2 \subset S$ are two subsets of an $\mathbb{F}$-universe, then the random variables $Y_i \colon S \to S/A_i$ are equivalent iff $A_1 = A_2$.*

Proof. Y_i partitions S into its A_i-cosets, one of which is A_i. Y_1 and Y_2 are equivalent iff they induce the same partition; since A_1, A_2 both contain the zero in S, if Y_1 and Y_2 are equivalent then $A_1 = A_2$. Conversely, if $A_1 = A_2$ then, $S \to S/A_i$ are the same map and hence equivalent. □

Recall that if $\mathcal{L} \colon V \to W$ is a linear map, then the map on dual spaces, $\mathcal{L}^* \colon W^* \to V^*$, has image equal to $\big(V/\ker(\mathcal{L})\big)^*$ viewed as "it sits" in V^*, i.e., viewed as the subspace of those elements of V^* that take $\ker(\mathcal{L})$ to zero. In particular, if $S \to S/A$ is a quotient map, then the image of the dual map is $(S/A)^*$ as it sits in S^*, i.e., the elements of S^* mapping all of A to zero (also referred to as the *annihilator* of A).

Recall that if S is any $\mathbb{F}$-universe and $A \subset S$ is a subspace, then the *annihilator of A in S^** is the set of elements of S^* taking all of A to 0, which is a subspace of dimension $\dim(S) - \dim(A)$; similarly, if $V \subset S^*$, by the *annihilator of V (in S)* we mean the elements of S that each element of V takes to zero, and that this is a subspace of dimension $\dim(S) - \dim(V)$.

Definition 2.8.4. Let $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, and let S be an $\mathbb{F}$-universe. By the *universe* associated to S we mean the dual space $\mathcal{U} = S^*$; by a *linear random variable* we mean a subspace $V \subset \mathcal{U}$, to which we associate the classical linear random variable $V_{\text{class}} \colon S \to S/A$ where A is the annihilator of V in S, i.e.,

$$A = \{s \in S \mid \forall \ell \in V, \ell(s) = 0\}.$$

Therefore, V equals the image of $(S/A)^*$ as it sits in S^*. We define the entropy of V to be that of V_{class}, i.e., $H_2(V) = H_2(V_{\text{class}})$.

From the above definition we can, conversely, associate any classical linear random variable to a unique linear random variable. For a classical linear random variable, $Y\colon S \to \mathcal{Y}$, consider the equivalent classical linear random variable $\tilde{Y}\colon S \to S/\ker(Y)$ which is a quotient map. Then the unique linear random variable associated to Y is $Y_{lrv} \subset S^*$ where Y_{lrv} is the image of $\big(S/\ker(Y)\big)^*$ as it sits in S^*, i.e., the annihilator of $\ker(Y)$ in S^*.

By above definition for a linear random variable, $V \subset S^*$, we have

$$H_2(V) = H_2(V_{\text{class}}) = \dim(S/A) = \dim\big((S/A)^*\big) = \dim(V).$$

Hence the entropy of V is its dimension. Similarly, the entropy of a classical linear random variable, Y, equals $\dim(Y_{lrv})$.

It will turn out to be far more convenient to think of a classical linear random variable as its associated linear random variable.

The last thing to note is how joint random variables work in the linear case. If Y_1, Y_2 are two random variables, then their joint random variable (Y_1, Y_2) denotes the random variable that is the Cartesian product map

$$Y_1 \times Y_2\colon S \to \mathcal{Y}_1 \times \mathcal{Y}_2.$$

If $S, \mathcal{Y}_1, \mathcal{Y}_2$ are vector spaces and Y_1, Y_2 are linear maps, then $\mathcal{Y}_1 \times \mathcal{Y}_2$ becomes a vector space and $Y_1 \times Y_2$ is a linear map.

Proposition 2.8.5. *Let $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, and let S be an $\mathbb{F}$-universe. Let $V^1, V^2 \subset \mathcal{U} = S^*$ be linear random variables, whose classical forms are $V^i_{\text{class}}\colon S \to S/A_i$ where A_i is the annihilator of V^i in S. Then the linear random variable associated to the classical random variable $(V^1_{\text{class}}, V^2_{\text{class}})$ is $V^1 + V^2$.*

Proof. Since

$$\ker\big((V^1_{\text{class}}, V^2_{\text{class}})\big) = \ker(V^1_{\text{class}} \times V^2_{\text{class}}) = \ker(V^1_{\text{class}}) \cap \ker(V^2_{\text{class}})$$

and V^i_{class} is a quotient map from S to S/A_i with kernel A_i we have

$$\ker\big((V^1_{\text{class}}, V^2_{\text{class}})\big) = A_1 \cap A_2.$$

Let V be the linear random variable associated to $(V^1_{\text{class}}, V^2_{\text{class}})$. Then V is the annihilator of $A_1 \cap A_2$ as it sits in S^* and has dimension equal to $\dim\big(S/(A_1 \cap A_2)\big)$. Since V^i is the annihilator of A_i, each V^i takes all of $A_1 \cap A_2$ to zero. Consequently, $(V^1 + V^2) \subset V$. By the dimension formula

$$\begin{aligned}\dim(V^1 + V^2) &= \dim(V^1) + \dim(V^2) - \dim(V^1 \cap V^2) \\ &= \dim(S/A_1) + \dim(S/A_2) - \dim(V^1 \cap V^2).\end{aligned}$$

Since $V^1 \cap V^2$ takes both A_1 and A_2 to zero it is the annihilator of $A_1 + A_2$ and has the same dimension as $S/(A_1 + A_2)$, then above equality becomes

$$\begin{aligned}\dim(V^1 + V^2) &= \dim(S/A_1) + \dim(S/A_2) - \dim\big(S/(A_1 + A_2)\big) \\ &= \dim(S) + \dim(A_1) + \dim(A_2) + \dim(A_1 + A_2) \\ &= \dim(S) - \dim(A_1 \cap A_2) = \dim\big(S/(A_1 \cap A_2)\big).\end{aligned}$$

Hence $V^1 + V^2$ and V have the same dimension. Since $V^1 + V^2 \subset V$, then $V = (V^1 + V^2)$. □

By above proposition for linear random variables, $V^1, V^2 \subset S^*$, we have

$$H_2(V^1_{class}, V^2_{class}) = H_2(V^1, V^2) = \dim(V^1 + V^2).$$

The chain rule of conditional entropy states for random variables X, Y

$$H(X \,|\, Y) = H(X, Y) - H(Y).$$

Using the chain rule we can define the conditional entropy of two linear random variables, $V^1, V^2 \subset S^*$, in terms of their corresponding classical linear random variables, V^1_{class}, V^2_{class}, as such

$$H_2(V^1 \,|\, V^2) = H_2(V^1_{class} \,|\, V^2_{class}) = H_2(V^1_{class}, V^2_{class}) - H_2(V^2_{class}).$$

It follows that

$$H_2(V^1 \,|\, V^2) = \dim(V^1 + V^2) - \dim(V^2),$$

using the dimension formula we get

$$H_2(V^1 \,|\, V^2) = \dim(V^1) - \dim(V^1 \cap V^2) = \dim^{\mathcal{U}/V^2}\left(\left[V^1\right]_{V^2}\right).$$

In the context of coded caching, for a random variable $Y\colon S \to \mathcal{Y}$, both S and $\mathcal{Y}$ are a $\mathbb{F}$-universe with $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$. So if Y is a classical linear random variable, then its corresponding linear random variable, $Y_{lrv} \subset S^*$, can be thought of as a linear subspace in S (since $S^* \cong S$ as S^* also has the structure of an $\mathbb{F}$-universe with the same dimension as S).

Chapter 3

Linear Information Theory

3.1 Coordination, Discoordination, and an Overview of the Main Results

In this section, we define the notion of the "discoordination" of a collection of subspaces of a universe, which is the focus of the linear algebra in this work. When a collection of subspaces have zero discoordination, when they are "coordinated," we get simple formulas regarding the dimensions of such subspaces and the dimension of subspaces obtained by applying $+, \cap$ operations and taking quotients.

The fact that two subspaces are always coordinated but three subspaces are not is known (see Exercise 9, Section 3.3 on page 51 of Jänich [6]). At the end of this section, we state the main theorem regarding the decomposition of any three subspaces of a universe into a coordinated part and a discoordinated part with a remarkably simple structure.

3.1.1 Coordination

If X is a set of linearly independent vectors in an $\mathbb{F}$-universe, $\mathcal{U}$, and $U \subset \mathcal{U}$ is a subspace, then $X \cap U$ is a set of linearly independent vectors in U, and hence

$$\dim(U) - |X \cap U| \geq 0 \tag{3.1.1}$$

with equality iff $X \cap U$ is a basis of U. This observation leads to a number of definitions that are the focus of this chapter.

Definition 3.1.1. Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. We say that a set of vectors, X, in $\mathcal{U}$ *coordinates* $U_1, \ldots, U_m$ if

1. X is a set of linearly independent vectors in $\mathcal{U}$, and
2. for all $i \in [m]$ we have
$$\dim(U_i) = |X \cap U_i|,$$
or, equivalently, $X \cap U_i$ is a basis for U_i (since $X \cap U_i \subset X$ is a linearly independent of vectors whose size equals the dimension of U_i).

If such an X exists, we say that $U_1, \ldots, U_m$ are *coordinated*.

Proposition 3.1.2. *If X coordinates subspaces U_1, U_2 of an $\mathbb{F}$-universe, $\mathcal{U}$, then X also coordinates $U_1 \cap U_2$ and $U_1 + U_2$.*

Proof. (An alternate proof is given in Subsection 3.1.2.) In view of (3.1.1), for any $X \in \mathrm{Ind}(\mathcal{U})$ we have

$$|X \cap (U_1 + U_2)| \leq \dim(U_1 + U_2), \quad |X \cap (U_1 \cap U_2)| \leq \dim(U_1 \cap U_2); \tag{3.1.2}$$

X coordinates $U_1 + U_2$ and $U_1 \cap U_2$ iff both these inequalities hold with equality, and the Inequality Summation Principle implies that equality holds in both iff their sum,

$$|X \cap (U_1 + U_2)| + |X \cap (U_1 \cap U_2)| \leq \dim(U_1 + U_2) + \dim(U_1 \cap U_2), \tag{3.1.3}$$

holds with equality. By (set theoretic) inclusion-exclusion we have

$$|X \cap (U_1 \cup U_2)| + |X \cap (U_1 \cap U_2)| = |X \cap U_1| + |X \cap U_2|,$$

since X coordinates U_1, U_2 we get that

$$|X \cap (U_1 \cup U_2)| + |X \cap (U_1 \cap U_2)| = \dim(U_1) + \dim(U_2).$$

From the dimension formula we know

$$\dim(U_1) + \dim(U_2) = \dim(U_1 + U_2) + \dim(U_1 \cap U_2),$$

therefore

$$|X \cap (U_1 \cup U_2)| + |X \cap (U_1 \cap U_2)| = \dim(U_1 + U_2) + \dim(U_1 \cap U_2). \tag{3.1.4}$$

However $U_1 \cup U_2$ is a subset of $U_1 + U_2$, and hence

$$|X \cap (U_1 + U_2)| + |X \cap (U_1 \cap U_2)| \geq |X \cap (U_1 \cup U_2)| + |X \cap (U_1 \cap U_2)|, \tag{3.1.5}$$

from (3.1.4) we have

$$|X \cap (U_1 + U_2)| + |X \cap (U_1 \cap U_2)| \geq \dim(U_1 + U_2) + \dim(U_1 \cap U_2).$$

This is the reverse inequality of (3.1.3), thus both hold with equality. □

We remark that (3.1.5) shows any element of $X \cap (U_1 + U_2)$ must also lie in $U_1 \cup U_2$.

It follows that if $\mathcal{U}$ is an $\mathbb{F}$-universe and $X \in \mathrm{Ind}(\mathcal{U})$ coordinates a family of subspaces, $\mathcal{W}$, then X also coordinates any subspace obtained by a finite sequence of spans and intersections of members of $\mathcal{W}$.

Another observation is if $\mathcal{U}$ is an $\mathbb{F}$-universe of dimension n, and $X \in \mathrm{Ind}(\mathcal{U})$, then X contains at most n vectors and hence X coordinates at most 2^n distinct subspaces of $\mathcal{U}$.

The essence of this chapter is to describe which subspaces $U_1, \ldots, U_m$ are coordinated, or, if not, to describe their "discoordination," which measures the extent to which they "fail to be coordinated." Before discussing discoordination, let us give a helpful way of thinking about coordinated subspaces.

3.1.2 Coordinate Subspaces

If $\mathbb{F}$ is a field, we use $\mathbb{F}^n$ to denote the usual product of n copies of $\mathbb{F}$, and use $e_1, \ldots, e_n$ to denote the standard basis vectors of $\mathbb{F}^n$ (hence e_i is a vector with a 1 in

the i-th coordinate and 0's elsewhere). For a subset $I \subset [n]$ where $I = \{i_1, \ldots, i_m\}$ for some $m \leq n$, we set e_I to be

$$e_I = \text{Span}(e_{i_1}, \ldots, e_{i_m}) \subset \mathbb{F}^n;$$

hence e_I is a subspace of dimension $|I| = m$ which we call the *I-coordinate subspace of* $\mathbb{F}^n$; consequently $e_\emptyset = \{0\}$ and $e_{[n]} = \mathbb{F}^n$. Note that if $I, J \subset [n]$, then

$$e_I + e_J = e_{I \cup J}, \quad e_I \cap e_J = e_{I \cap J}. \tag{3.1.6}$$

This gives us another view of coordination: if $\mathcal{U}$ is an n-dimensional $\mathbb{F}$-universe and $X = \{x_1, \ldots, x_n\}$ is a basis of $\mathcal{U}$, then there is a unique isomorphism $f : \mathcal{U} \to \mathbb{F}^n$ of vector spaces such that $x_i \in \mathcal{U}$ is taken to $e_i \in \mathbb{F}^n$. In this case a subspace $W \subset \mathcal{U}$ is coordinated by X iff $f(W)$ is a coordinate subspace in $\mathbb{F}^n$.

We can reprove Proposition 3.1.2 from this perspective: let $U_1, U_2 \subset \mathcal{U}$ be coordinated by $X \in \text{Ind}(\mathcal{U})$; we may extend X to be a basis on all of $\mathcal{U}$, which still coordinates U_1, U_2. Now we have an isomorphism $f : \mathcal{U} \to \mathbb{F}^n$; since an isomorphism of vector spaces preserves the operations $+, \cap$, we have $f(U_1) = e_I$ where $I \subset [m]$ consists of those $i \in [m]$ such that $e_i \in f(U_1 \cap X)$; similarly define $J \subset [m]$ such that $f(U_2) = e_J$. In view of (3.1.6), and the fact that $f^{-1} : \mathbb{F}^n \to \mathcal{U}$ is another isomorphism, preserving $+$ and $\cap$, we have that

$$U_1 + U_2 = f^{-1}(e_{I \cup J}), \quad U_1 \cap U_2 = f^{-1}(I \cap J)$$

are coordinated by $f^{-1}(e_1), \ldots, f^{-1}(e_n)$, which are the elements of X.

3.1.3 Discoordination and Minimizers

In this section we define a measure for "the extent to which given subspaces of a universe may fail to be coordinated."

Definition 3.1.3. Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. If $X \in \text{Ind}(\mathcal{U})$ (i.e., X is a subset of linearly independent vectors in $\mathcal{U}$), we define the *discoordi-*

nation of $U_1, \ldots, U_m$ with respect to X to be

$$\mathrm{DisCoord}_X(U_1, \ldots, U_m) = \sum_{i=1}^{m} \Big(\dim(U_i) - |X \cap U_i| \Big).$$

We define the *discoordination of $U_1, \ldots, U_m$* to be

$$\mathrm{DisCoord}(U_1, \ldots, U_m) = \min_{X \in \mathrm{Ind}(\mathcal{U})} \mathrm{DisCoord}_X(U_1, \ldots, U_m),$$

we call any $X \in \mathrm{Ind}(\mathcal{U})$ at which the above minimum is attained a *minimizer of* $U_1, \ldots, U_m$.

In view of (3.1.1), $U_1, \ldots, U_m$ are coordinated iff their discoordination equals 0, and, if so, then X is a minimizer of $U_1, \ldots, U_m$ when X coordinates $U_1, \ldots, U_m$.

Note that in the above definition, if $X \subset X'$ and $X' \in \mathrm{Ind}(\mathcal{U})$, then

$$\mathrm{DisCoord}_X(U_1, \ldots, U_m) \geq \mathrm{DisCoord}_{X'}(U_1, \ldots, U_m).$$

It follows that if X is a minimizer of $U_1, \ldots, U_m$ and not a basis of $\mathcal{U}$, we can extend X to obtain a basis X' of $\mathcal{U}$ such that $X \subset X'$, where X' is also a minimizer of $U_1, \ldots, U_m$. Hence there are always minimizers that are bases for $\mathcal{U}$.

3.1.4 The Main Theorem Regarding Three Subspaces

Theorem 3.1.4. *Let U_1, U_2, U_3 be three subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$ and let $S_2 = (U_1 \cap U_2) + (U_1 \cap U_3) + (U_2 \cap U_3)$. Then the following non-negative integers are equal:*

1. $\mathrm{DisCoord}(U_1, U_2, U_3)$*;*
2. *the minimum of* $\dim(U_3) - |X \cap U_3|$ *over all* $X \in \mathrm{Ind}(\mathcal{U})$ *that coordinate* U_1 *and* U_2*;*
3. *the dimension in* $\mathcal{U}/S_2$ *of the space* $([U_1]_{S_2} + [U_2]_{S_2}) \cap [U_3]_{S_2}$.

Furthermore, if $m = \mathrm{DisCoord}(U_1, U_2, U_3)$*, then there is a decomposition* $\mathcal{U}_1, \mathcal{U}_2$ *of* $\mathcal{U}$ *through which* U_1, U_2, U_3 *all factor,*

1. *$U_1 \cap \mathcal{U}_1, U_2 \cap \mathcal{U}_1, U_3 \cap \mathcal{U}_1$ are coordinated in $\mathcal{U}_1$, and*

2. *there is an isomorphism $\mu : \mathcal{U}_2 \to \mathbb{F}^2 \times \mathbb{F}^m$ which takes $U_1 \cap \mathcal{U}_2, U_2 \cap \mathcal{U}_2, U_3 \cap \mathcal{U}_2$, respectively to*

$$\mathrm{Span}(e_1) \otimes \mathbb{F}^m,\ \mathrm{Span}(e_2) \otimes \mathbb{F}^m,\ \mathrm{Span}(e_1 + e_2) \otimes \mathbb{F}^m.$$

After proving this theorem (proof given in Section 3.4), we will be able to write a number of important formulas involving U_1, U_2, U_3 in terms of their discoordination. Let us first state the general principle.

Definition 3.1.5. Let $f = f(U_1, \ldots, U_m)$ be a formula that is an $\mathbb{Z}$-linear combination of terms of the form $\dim(U'/U'')$, where U', U'' are formulas in the operations $\cap, +$ and the variables $U_1, \ldots, U_m$ (and parenthesis); hence f takes arbitrary subspaces $U_1, \ldots, U_m$ of some $\mathbb{F}$-universe, $\mathcal{U}$, and returns an integer. We say that f is a *balanced formula* if $f(U_1, \ldots, U_m) = 0$ whenever $U_1, \ldots, U_m$ are coordinated.

Corollary 3.1.6. *Let $f = f(U_1, U_2, U_3)$ be a balanced formula. Then for any subspaces $U_1, U_2, U_3 \subset \mathcal{U}$ of a $\mathbb{F}$-universe, $\mathcal{U}$, we have*

$$f(U_1, U_2, U_3) = k\ \mathrm{DisCoord}(U_1, U_2, U_3),$$

where

$$k = f\big(\mathrm{Span}(e_1), \mathrm{Span}(e_2), \mathrm{Span}(e_1 + e_2)\big),$$

and e_1, e_2 are the standard basis vectors in $\mathbb{F}^2$ for any field $\mathbb{F}$.

The above corollary is an immediate consequence of Theorem 3.1.4 and Theorem 2.5.7 and the paragraph just below it, since

$$f(U_1, U_2, U_3) = \sum_{i=1}^{2} f(U_1 \cap \mathcal{U}_i, U_2 \cap \mathcal{U}_i, U_3 \cap \mathcal{U}_i),$$

where $\mathcal{U}_1$,$\mathcal{U}_2$ are the decomposition of $\mathcal{U}$ stated Theorem 3.1.4; thus in the equality above the $i = 1$ term vanishes since this term involves coordinated subspaces, and the $i = 2$ term is isomorphic to the direct sum of m copies of $\mathbb{F}^2$, in which f restricted to each copy equals k above.

Corollary 3.1.7. *Let U_1, U_2, U_3 be three subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$. Then* $\mathrm{DisCoord}(U_1, U_2, U_3)$ *equals*

1. $m = \dim(U_1 \cap U_2 \cap U_3) - I(U_1; U_2; U_3)$, *where*

$$\begin{aligned} I(U_1; U_2; U_3) = &\dim(U_1 + U_2 + U_3) \\ &- \dim(U_1 + U_2) - \dim(U_1 + U_3) - \dim(U_2 + U_3) \\ &+ \dim(U_1) + \dim(U_2) + \dim(U_3). \end{aligned}$$

2. $\dim\big(U_3 \cap (U_1 + U_2)\big) - \dim(U_3 \cap U_1) - \dim(U_3 \cap U_2) + \dim(U_1 \cap U_2 \cap U_3)$.

3. $\dim\big((U_1 + U_3) \cap (U_2 + U_3)\big) - \dim(U_3) + \dim(U_1 \cap U_2 \cap U_3) - \dim(U_1 \cap U_2)$.

Corollary 3.1.7 will be shown in Section 3.6, specifically we will show all formulas in the corollary are balanced. In Subsection 3.6.1 we will show other equalities involving the discoordination of three subspaces in some quotient space.

3.1.5 Discoordination in Quotient Spaces

Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. For a subspace $W \subset \mathcal{U}$, it is interesting to consider the relationship between $\mathrm{DisCoord}(U_1, \ldots, U_m)$ in $\mathcal{U}$, and the discoordination of the images of $U_1, \ldots, U_m$ in the quotient universe $\mathcal{U}/W$,

$$\mathrm{DisCoord}^{\mathcal{U}/W}\big([U_1]_W, \ldots, [U_m]_W\big)$$

(where we use the superscript $\mathcal{U}/W$ to emphasize that we consider the $[U_i]_W = U_i + W$ as subspaces of the quotient space $\mathcal{U}/W$).

Let us give examples that show that this new discoordination, in "passing from $\mathcal{U}$ to $\mathcal{U}/W$," can decrease or increase.

If $W = \mathcal{U}$, then $\mathcal{U}/W$ is the zero dimensional universe, where the discoordination is always zero; suppose $U_1, \ldots, U_m$ are not coordinated (e.g., $m = 3$ and U_1, U_2, U_3 are, respectively, the spans of $e_1, e_2, e_1 + e_2$ in $\mathbb{F}^2$), then the discoordination can decrease in passing from $\mathcal{U}$ to $\mathcal{U}/W$ (in this example the discoordination decreases from 1 in $\mathcal{U}$ to 0 in $\mathcal{U}/W$).

On the other hand, if U_1, U_2, U_3 are, respectively, the spans of e_1, e_2, e_3 in $\mathbb{F}^3$, then they are coordinated. However in $\mathcal{U}/W$ with $W = \mathrm{Span}(e_1 + e_2 - e_3)$, we have

that $[e_1]_W$ and $[e_2]_W$ are, respectively, bases for $[U_1]_W$ and $[U_2]_W$, while also forming a basis for the quotient space $\mathcal{U}/W$; additionally, $[e_1+e_2]_W$ is a basis for $[U_3]_W$, this means $[U_1]_W, [U_2]_W, [U_3]_W$ form the discoordination example from (1.2.1) and cannot be coordinated. Hence U_1, U_2, U_3 can be coordinated in $\mathcal{U}$ but have positive discoordination in $\mathcal{U}/W$.

Theorem 3.1.4 already implies[1] that for $S_2 = (U_1 \cap U_2) + (U_1 \cap U_3) + (U_2 \cap U_3)$,

$$\text{DisCoord}^{\mathcal{U}}(U_1, U_2, U_3) = \text{DisCoord}^{\mathcal{U}/S_2}\big([U_1]_{S_2}, [U_2]_{S_2}, [U_3]_{S_2}\big).$$

In Theorems 3.2.10 and 3.2.11 we show a more general result related to the equality above for arbitrarily many subspaces. The following is another useful property of the discoordination of three subspaces.

Theorem 3.1.8. *Let U_1, U_2, U_3, W be four subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$, such that $W \subset (U_1 \cap U_2)$. Then,*

$$\text{DisCoord}^{\mathcal{U}}(U_1, U_2, U_3) = \text{DisCoord}^{\mathcal{U}/W}\big([U_1]_W, [U_2]_W, [U_3]_W\big).$$

The discoordination of U_1, U_2, U_3 in $\mathcal{U}$ is the same as that of the images of U_1, U_2, U_3 in the quotient $\mathcal{U}/W$.

A proof for Theorem 3.1.8 is given in Section 3.5.

3.1.6 The Discoordination Formula

Some of the results in this work are based on a detailed description of how to build minimizers for subspaces $U_1, \ldots, U_m$ in a universe. This description gives an interesting "formula" for discoordination which we will use to prove Theorem 3.1.8. Both results are stated as a single theorem, namely Theorem 3.2.9. In an effort to present the important results at the beginning and not loose sight of them in the sea of proofs that follow, let us separately state the discoordination formula.

Theorem 3.1.9. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. For each $k \in [m]$,*

[1] This equality is also implied by Theorem 3.2.11 and the full statement of Theorem 3.1.4 is not necessary for the implication.

let S_k be the span of all intersections of any k of $U_1,\dots,U_m$, i.e.,

$$S_k = \mathcal{S}_k(U_1,\dots,U_m) = \sum_{1\le i_1<\dots<i_k\le m} U_{i_1}\cap\dots\cap U_{i_k}$$

(see Definition 3.2.6). Then

$$\mathrm{DisCoord}(U_1,\dots,U_m) = \sum_{i=1}^{m}\dim(U_i) - \sum_{i=1}^{m}\dim(S_i).$$

What makes this theorem difficult to apply is that one needs information about $S_2,\dots,S_m$, which is detailed information based on $U_1,\dots,U_m$; usually it is only $S_1 = U_1+\dots+U_m$ that tends to be clear in applications.

3.1.7 Factorization and Discoordination

Let us describe an application of Theorem 3.1.9 based on factorization. This will lead into another important property of discoordination. By Theorem 2.5.7 and Theorem 3.1.9 we have the following important equality.

Theorem 3.1.10. *Let $A_1,\dots,A_m$ be subspaces of an $\mathbb{F}$-universe $\mathcal{U}$ that all factor through a decomposition $\mathcal{U}_1,\dots,\mathcal{U}_r$ of $\mathcal{U}$. Then*

$$\mathrm{DisCoord}^{\mathcal{U}}(A_1,\dots,A_m) = \sum_{i=1}^{r}\mathrm{DisCoord}^{\mathcal{U}_i}\big(A_1\cap\mathcal{U}_i,\dots,A_m\cap\mathcal{U}_i\big).$$

Proof. Theorem 2.5.7 implies that the subspaces $S_1,\dots,S_m$ of Theorem 3.1.9 can be computed as the $S_1,\dots,S_m$ of $A_1\cap\mathcal{U}_i,\dots,A_m\cap\mathcal{U}_i$ and summed.

Formally, by Theorem 3.1.9 we have

$$\mathrm{DisCoord}(A_1,\dots,A_m) = \sum_{k=1}^{m}\big(\dim(A_k)-\dim(S_k)\big),$$

where

$$S_k = \mathcal{S}_k(A_1,\dots,A_m) = \sum_{1\le i_1<\dots<i_k\le m} A_{i_1}\cap\dots\cap A_{i_k}.$$

Since the A_k factor through the decomposition, for $k \in [m]$

$$\dim(A_k) = \sum_{i=1}^{r} \dim(A_k \cap \mathcal{U}_i).$$

By Theorem 2.5.7 the S_k also factor through the decomposition and

$$S_k = \sum_{1 \le i_1 < \ldots < i_k \le m} \sum_{j=1}^{r} \big((A_{i_1} \cap \ldots \cap A_{i_k}) \cap \mathcal{U}_j \big),$$

which implies

$$S_k = \sum_{i=1}^{r} \mathcal{S}_k\big((A_1 \cap \mathcal{U}_i), \ldots, (A_m \cap \mathcal{U}_i) \big).$$

Hence we can write

$$\text{DisCoord}(A_1, \ldots, A_m) = \sum_{i=1}^{r} \sum_{k=1}^{m} \Big(\dim(A_k \cap \mathcal{U}_i) - \mathcal{S}_k\big((A_1 \cap \mathcal{U}_i), \ldots, (A_m \cap \mathcal{U}_i) \big) \Big),$$

which simplifies to the theorem hypothesis. □

3.1.8 Additional Results about Coordination

Another noteworthy, yet unused, theorem in coordination (the results in coded caching we present in Chapter 4 circumvent the need for it) is Theorem 3.3.1. Theorem 3.3.1 states that if $A_1 \subset \cdots \subset A_s$ and $B_1 \subset \cdots \subset B_t$ are two increasing sequences of subspaces of an $\mathbb{F}$-universe, then all these subspaces and their pairwise intersections $\big($i.e. $A_i \cap B_j$ for any $i \in [s]$ and $j \in [t]\big)$ are coordinated.

In particular, Theorem 3.3.1 implies the sometimes convenient fact that if A, B, C are subspaces of an $\mathbb{F}$-universe with $B \subset C$, then A, B, C are coordinated.

3.2 Minimizers and Greedy Algorithms

In this section we prove theorems regarding the structure of minimizers and drive the discoordination "formula" from Theorem 3.1.9. As previously stated, without a good understanding of the S_i we only get partial information about the discoordination. Still, this discoordination formula, and certain other considerations based on "greedy algorithms" to build minimizers, will make proving Theorem 3.1.4 easier.

3.2.1 Meet Numbers and Basic Greedy Considerations

There are numerous properties of discoordination minimizers, X, of subsets $U_1,\dots,U_m$ worth stating. Since

$$\mathrm{DisCoord}_X(U_1,\dots,U_m)=\sum_{i=1}^{m}\Big(\dim(U_i)-|X\cap U_i|\Big)=\sum_{i=1}^{m}\dim(U_i)-\sum_{i=1}^{m}|X\cap U_i|,$$

a minimizer X is an element of $\mathrm{Ind}(\mathcal{U})$ that maximizes $\sum_{i=1}^{m}|X\cap U_i|$.

Definition 3.2.1. Let $U_1,\dots,U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. For a finite subset $X\subset\mathcal{U}$ we define the *meet of X in $U_1,\dots,U_m$* to be

$$\mathrm{Meet}(X)=\mathrm{Meet}(X;U_1,\dots,U_m)=\sum_{i=1}^{m}|X\cap U_i|.$$

If $x\in\mathcal{U}$, we define the *(point-wise) meet of x in $U_1,\dots,U_m$* to be

$$\mathrm{meet}(x)=\mathrm{meet}(x;U_1,\dots,U_m)=\mathrm{Meet}(\{x\};U_1,\dots,U_m)=\big|\{i\in[m]\mid x\in U_i\}\big|.$$

If $X=\{x_1,\dots,x_m\}$ is a finite subset of $\mathcal{U}$, we say that $x_1,\dots,x_m$ is in *decreasing meet order* if

$$\mathrm{meet}(x_1)\geq\mathrm{meet}(x_2)\geq\cdots\geq\mathrm{meet}(x_n).$$

Usually $U_1,\dots,U_m$ will be fixed, so we may simply write $\mathrm{Meet}(X)$ and $\mathrm{meet}(x)$ without confusion. Of course $\mathrm{meet}(x)$ is the same as $\mathrm{Meet}(\{x\})$; hence we distinguish between "meet" and "Meet" for ease of reading.

Given the above definition we have

$$\mathrm{Meet}(X;U_1,\dots,U_m)=\sum_{x\in X}\mathrm{meet}(x;U_1,\dots,U_m).$$

It follows that $X\in\mathrm{Ind}(\mathcal{U})$ that minimize

$$\mathrm{DisCoord}_X(U_1,\dots,U_m)=\sum_{i=1}^{m}\dim(U_i)-\mathrm{Meet}(X;U_1,\dots,U_m),$$

are the same $X \in \mathrm{Ind}(\mathcal{U})$ that maximize

$$\mathrm{Meet}(X; U_1, \ldots, U_m) = \sum_{i=1}^{m} |X \cap U_i| = \sum_{x \in X} \mathrm{meet}(x; U_1, \ldots, U_m).$$

The following is an important remark about minimizers that is related to the "greedy algorithm" we will discuss in Subsection 3.2.2.

Proposition 3.2.2. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let $X \in \mathrm{Ind}(\mathcal{U})$ be a minimizer of $U_1, \ldots, U_m$, which is a basis of $\mathcal{U}$ and is arranged in a meet decreasing order, i.e., $X = \{x_1, \ldots, x_n\}$, and*

$$\mathrm{meet}(x_1) \geq \mathrm{meet}(x_2) \geq \cdots \geq \mathrm{meet}(x_n).$$

Let $X' = \{x'_1, \ldots, x'_k\}$ be any other independent set in $\mathcal{U}$ arranged in meet decreasing order, i.e.,

$$\mathrm{meet}(x'_1) \geq \mathrm{meet}(x'_2) \geq \cdots \geq \mathrm{meet}(x'_k).$$

Then $\mathrm{meet}(x_k) \geq \mathrm{meet}(x'_k)$.

Proof. We claim that for some $j' \in [k]$ and some $j \in \{k, k+1, \ldots, n\}$, we may exchange $x'_{j'}$ for x_j in X and get a new independent set of vectors $X'' = (X \setminus \{x_j\}) \cup \{x'_{j'}\}$. To show this, consider that each $x'_{j'}$ may be written uniquely as a linear combination

$$\gamma_{j'1} x_1 + \gamma_{j'2} x_2 + \cdots + \gamma_{j'n} x_n$$

(where the $\gamma_{j'i} \in \mathbb{F}$ are scalars). If $\gamma_{j'i} \neq 0$, in X we may exchange x_i with $x'_{j'}$ and get a new basis. Hence we do this for some $i \geq k$ unless $\gamma_{j'i} = 0$ for all $j' \in [k]$ and $i \geq k$, which implies that for all $j' \in [k]$,

$$x'_{j'} \in W = \mathrm{Span}(x_1, \ldots, x_{k-1}).$$

However, $x'_1, \ldots, x'_k$ is a set of k linearly independent vectors, and cannot all lie in the $(k-1)$-dimensional subspace W. Therefore, there is some $x'_{j'}$ with $j' \leq k$ that can be substituted for an $x_i \in X$ with $i \geq k$ to obtain a new basis X''.

Suppose for the sake of contradiction that $\text{meet}(x_k) < \text{meet}(x'_k)$ then

$$\text{meet}(x_i) \leq \text{meet}(x_k) < \text{meet}(x'_k) \leq \text{meet}(x'_{j'}),$$

and hence replacing x_i in the basis X with $x'_{j'}$ gives a basis X'' with $\text{Meet}(X'') > \text{Meet}(X)$. This contradicts the discoordination minimality of X. □

The above proposition implies that if X, X' are two minimizers of $U_1, \ldots, U_m$, both arranged in a decreasing meet order $x_1, \ldots, x_n$ and $x'_1, \ldots, x'_{n'}$, then $\text{meet}(x_i) = \text{meet}(x'_i)$ for $i \in [\min(n, n')]$.

There are a few important corollaries of the above proposition on substitution of vectors in a minimizing set of linearly independent vectors that we must emphasise.

Theorem 3.2.3. *Let $X \in \text{Ind}(\mathcal{U})$ be a minimizer of subsets $U_1, \ldots, U_m$ of some $\mathbb{F}$-universe, $\mathcal{U}$, and let $X' = X \cap (U_1 + \cdots + U_m)$. Then the following are true*

1. $|X'| = \dim(U_1 + \cdots + U_m)$
2. *if $x' \in X'$, then* $\text{meet}(x') \geq 1$
3. *if $x \in X \setminus X'$, then* $\text{meet}(x) = 0$

Proof. Since $W = U_1 + \cdots + U_m$ is spanned by $U_1 \cup \cdots \cup U_m$, one can write W as the span of $\dim(W)$ vectors, each of which lies in at least one U_i. If $n = \dim(W)$, this gives n linearly independent vectors $x_1, \ldots, x_n$ for which $\text{meet}(x_i) \geq 1$ for all $i \in [n]$. It follows that X has at least n vectors, x, that have $\text{meet}(x) \geq 1$. Since these n vectors all lie in X', and are linearly independent, they form a basis for W, and all other vectors in X must lie outside $U_1 + \cdots + U_m$. □

Definition 3.2.4. Let $X \in \text{Ind}(\mathcal{U})$ be a minimizer of subsets $U_1, \ldots, U_m$ of some $\mathbb{F}$-universe, $\mathcal{U}$. We say that X is a *small minimizer* (with respect to $U_1, \ldots, U_n$) if $X \subset (U_1 + \ldots + U_m)$, and is a *large minimizer* if X is a basis for $\mathcal{U}$.

Proposition 3.2.5. *Let $X \in \text{Ind}(\mathcal{U})$ be a minimizer of subsets $U_1, \ldots, U_m$ of some $\mathbb{F}$-universe, $\mathcal{U}$. Then $X' \subset X \subset X''$ where X' is a small minimizer and X'' is a large minimizer. Furthermore all small minimizers are of size* $\dim(U_1 + \cdots + U_m)$.

Proof. We have $X' = X \cap (U_1 + \ldots + U_m)$ is a small minimizer and $X' \subset X$. By Theorem 3.2.3, X' is of size $\dim(U_1 + \cdots + U_m)$ and spans all of $U_1 + \cdots + U_m$. If X is not a basis for $\mathcal{U}$ we can extend it to a basis X'' of $\mathcal{U}$. It follows that all elements of $X'' \setminus X'$ lie outside of $U_1 + \cdots + U_m$, and hence each has meet zero with $U_1, \ldots, U_m$. □

3.2.2 The Greedy Algorithm for Minimizers

We give a simple "greedy algorithm" to build a minimizer, X, of subspaces $U_1, \ldots, U_m$ of a universe. This algorithm will help us prove Theorems 3.1.4, 3.1.8, and 3.1.9.

Definition 3.2.6. Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. For any $k \in [m]$, a *k-fold intersection of the* $U_1, \ldots, U_m$ refers to any subspace of the form

$$U_{i_1} \cap \ldots \cap U_{i_k} \quad \text{where} \quad 1 \leq i_1 < \ldots < i_k \leq m.$$

For each $k \in [m]$, we use S_k, V_k respectively to denote the respective sum and union of all the k-fold intersections of the $U_1, \ldots, U_m$, i.e.,

$$S_k = \mathcal{S}_k(U_1, \ldots, U_m) = \sum_{1 \leq i_1 < \ldots < i_k \leq m} U_{i_1} \cap \ldots \cap U_{i_k},$$
$$V_k = \mathcal{V}_k(U_1, \ldots, U_m) = \bigcup_{1 \leq i_1 < \ldots < i_k \leq m} U_{i_1} \cap \ldots \cap U_{i_k},$$

where the $\mathcal{S}_k$ and the $\mathcal{V}_k$ are functions defining the S_k and the W_k in terms of $U_1, \ldots, U_m$. For instance we have,

$$S_1 = U_1 + \cdots + U_m, \quad S_2 = \sum_{i_1 < i_2} U_{i_1} \cap U_{i_2},$$

$$S_3 = \sum_{i_1 < i_2 < i_3} U_{i_1} \cap U_{i_2} \cap U_{i_3}, \quad S_m = U_1 \cap \ldots \cap U_m;$$

and

$$V_1 = U_1 \cup \cdots \cup U_m, \quad V_2 = \bigcup_{i_1 < i_2} U_{i_1} \cap U_{i_2},$$

$$V_3 = \bigcup_{i_1<i_2<i_3} U_{i_1} \cap U_{i_2} \cap U_{i_3}, \quad V_m = U_1 \cap \ldots \cap U_m.$$

For convenience[2] we set $S_{m+1} = 0$, $V_{m+1} = \emptyset$, and $S_0 = V_0 = \mathcal{U}$.

Recall from Subsection 3.2.1 that $X \in \mathrm{Ind}(\mathcal{U})$ is a minimizer of $U_1, \ldots, U_m$ iff it maximizes the expression

$$f(X) = \sum_{x \in X} \mathrm{meet}(x) = \sum_{x \in X} \mathrm{meet}(x; U_1, \ldots, U_m). \tag{3.2.1}$$

It is important to note that the V_i defined above are subsets of $\mathcal{U}$ and not generally subspaces of $\mathcal{U}$, however the S_i are subspaces of $\mathcal{U}$. The V_i and S_i satisfy the following conditions:

1. for all $i = 0, \ldots, m$ we have $V_i \subset S_i = \mathrm{Span}(V_i)$;
2. $0 = S_{m+1} \subset S_m \subset \ldots \subset S_1 \subset S_0 = \mathcal{U}$;
3. $\emptyset = V_{m+1} \subset V_m \subset \ldots \subset V_1 \subset V_0 = \mathcal{U}$;
4. for all $y \in \mathcal{U}$ and $i = 0, \ldots, m$ we have $\mathrm{meet}(y) = i$ iff $y \in V_i \setminus V_{i+1}$;
5. for all $i = 0, \ldots, m$ we have that S_i/S_{i+1} is spanned by the images of the elements of V_i in the quotient space $\mathcal{U}/S_{i+1}$; hence
6. for all $i = 0, \ldots, m$, there is a subset $Y_i \subset V_i$ whose image in $\mathcal{U}/S_{i+1}$ is a basis for S_i/S_{i+1}; since no such element of Y_i can lie in S_{i+1} (i.e., equal 0 in $\mathcal{U}/S_{i+1}$), we have that any such Y_i consists entirely of elements $y \in \mathcal{U}$ such that $\mathrm{Meet}(y) = i$.

The Y_i described above turn out to be essential to our greedy algorithm.

Definition 3.2.7. Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. For $i = 0, \ldots, m$, we say that a set Y_i is a *purely i-th intersection basis (for $U_1, \ldots, U_m$)* if

1. Y_i is a basis of S_i relative to S_{i+1}, and

[2] Our conveniences actually follow by definition, given reasonable interpretations of an empty intersection, empty sum, and an empty union.

2. for any $y \in Y_i$, $\text{meet}(y) = \text{meet}(y; U_1, \ldots, U_m) = i$.

Proposition 3.2.8. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. Then for each $i = 0, \ldots, m$, there exists a purely i-th intersection basis.*

Proof. The proof consists of unwinding the definitions. Setting $V_i' = V_i \setminus V_{i+1}$ we have

$$V_i' = \{y \in \mathcal{U} \mid \text{meet}(y; U_1, \ldots, U_m) = i\}.$$

Since $V_i = V_i' \cup V_{i+1}$, $S_i = \text{Span}(V_i)$, and $S_{i+1} = \text{Span}(V_{i+1})$, it follows that S_i is spanned by S_{i+1} and V_i'. Hence, there exists a Y_i consisting entirely of elements of V_i' such that Y_i is a basis of S_i relative to S_{i+1}. □

Roughly, our "greedy algorithm" to construct a minimizer, X, of subspaces $U_1, \ldots, U_m$ of a universe functions as such:

- since a minimizer $X \in \text{Ind}(\mathcal{U})$ maximizes the expression in (3.2.1), our "greedy algorithm" first chooses the largest possible subset $Y_m \in \text{Ind}(\mathcal{U})$ consisting of elements in y with $\text{Meet}(y) = m$; hence Y_m can be as large as $\dim(S_m)$, and such a Y_m is a basis for S_m;
- the second step is choose the largest subset, Y_{m-1}, possible consisting of y with $\text{meet}(y) = m-1$ such that $Y_m \cup Y_{m-1}$ remains linearly independent; the largest possible Y_{m-1} is of size $\dim(S_{m-1}/S_m)$ and must be a purely $(m-1)$-th intersection basis (evident from Theorem 3.2.9);
- the i-th step, for $i = 3, \ldots, m$ is that given $Y_m, Y_{m-1}, \ldots, Y_{m-i+2}$, we choose Y_{m-i+1} to consist of $y \in \mathcal{U}$ with $\text{meet}(y) = m - i + 1$ and as large as possible with $Y_m \cup Y_{m-1} \cup \cdots \cup Y_{m-i+1}$ linearly independent.

We will prove that the algorithm roughly described above always produces a minimizer, and each minimizer is constructed as such. A novel point is that each Y_i is independent of the particularly chosen $Y_m, \ldots, Y_{i+1}$, since one can take Y_i to be an arbitrary purely i-th intersection basis. Let us state this result formally.

Theorem 3.2.9. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. Let $X \in \mathrm{Ind}(\mathcal{U})$, and for $i = 0, \ldots, m$ set*

$$Y_i = \big\{ x \in X \mid \mathrm{meet}(x; U_1, \ldots, U_m) = i \big\}.$$

Then

$$\mathrm{Meet}(X; U_1, \ldots, U_m) \leq \sum_{i=1}^{m} |S_i|, \tag{3.2.2}$$

with equality iff for each $i \in [m]$, Y_i is a purely i-intersection basis, i.e., Y_i is a basis of S_i relative to S_{i+1} and all elements of Y_i lie in exactly i of $U_1, \ldots, U_m$. Such Y_i exist by Proposition 3.2.8.

Moreover, all minimizers, $X \in \mathrm{Ind}(\mathcal{U})$, of $U_1, \ldots, U_m$ are of the form

$$X = Y_m \cup \cdots \cup Y_1 \cup Y_0,$$

where for $i \in [m]$, Y_i is a purely i-intersection basis, and Y_0 is some set of vectors that is a subset of a purely 0-intersection basis, i.e., a set of vectors whose image in $S_0/S_1 = \mathcal{U}/S_1$ is an independent set in $\mathcal{U}/S_1$. In particular, for any such X we have

$$|Y_i| + \cdots + |Y_m| = \dim(S_i), \tag{3.2.3}$$

and

$$Y_i, \ldots, Y_m \text{ are mutually disjoint and are a basis of } S_i \text{ for all } i \geq 1. \tag{3.2.4}$$

We also have

$$\sum_{i=1}^{m} \dim(S_i) = \sum_{i=1}^{m} i\,|Y_i| = \sum_{i=1}^{m} i \dim(S_i/S_{i+1}), \tag{3.2.5}$$

and hence we may also write

$$\mathrm{DisCoord}(U_1, \ldots, U_m) = \sum_{i=1}^{m} \dim(U_i) - \sum_{i=1}^{m} i \dim(S_i/S_{i+1}) \tag{3.2.6}$$

Finally

$$\mathrm{DisCoord}(U_1, \ldots, U_m) = \sum_{i=1}^{m} \dim(U_i) - \sum_{i=1}^{m} \dim(S_i). \tag{3.2.7}$$

Proof. By definition, the $Y_0, \ldots, Y_m$ are disjoint, since if $y \in Y_i$ then $\text{meet}(y) = i$ for all $i = 0, 1, \ldots, m$. For any $i \in [m]$ since $Y_i \cup \cdots \cup Y_m$ lies in S_i and is a subset of X, these vectors are linearly independent, and hence

$$|Y_i| + \cdots + |Y_m| \leq \dim(S_i). \tag{3.2.8}$$

Using the inequality summation principle, we can sum (3.2.8) for all $i \in [m]$ to first obtain the inequality

$$\sum_{i=1}^{m} i\,|Y_i| \leq \sum_{i=1}^{m} \dim(S_i), \tag{3.2.9}$$

and infer that equality holds in (3.2.9) iff for all $i \in [m]$ we have (3.2.8) holding with equality. Since

$$\text{Meet}(X; U_1, \ldots, U_m) = \sum_{i=1} \text{Meet}(Y_i; U_1, \ldots, U_m) = \sum_{i=1}^{m} i\,|Y_i|,$$

we can rewrite (3.2.9) as (3.2.2).

Next consider when (3.2.2) holds with equality. Accordingly, (3.2.8) must hold with equality for each $i \in [m]$. For $i = m$ we must have $|Y_m| = \dim(S_m)$, this happens iff Y_m is a basis for S_m; since $S_{m+1} = 0$, this also implies that Y_m is a basis for S_m relative to S_{m+1}.

For $i = m-1$, (3.2.9) holds iff $Y_m \cup Y_{m-1}$ is a basis for S_{m-1}; as Y_m is a basis for S_m, this holds with equality iff Y_{m-1} is a basis for S_{m-1} relative to S_m. Since Y_{m-1} are the elements of X that meet exactly $m-1$ of $U_1, \ldots, U_m$, such Y_{m-1} exist and are a purely $(m-1)$-intersection bases.

We proceed by investigating (3.2.9) for $j = m-2$, then $j = m-3$, and so on until $j = 1$. For each such j, we have established that (3.2.9) holds for all $i \geq j+1$ exactly when for each $i \geq j+1$, Y_i is a purely i-th intersection basis. Hence $Y_m \cup \cdots \cup Y_{j+1}$ is a basis for S_{j+1}, and the same argument as in the previous paragraph shows that (3.2.9) holds with equality for $i = j$ iff Y_j is a purely j-th intersection basis.

It follows that there exist X satisfying (3.2.2), and all such X have each Y_i being a purely i-th intersection basis for $i \geq 1$. Y_0 has no effect on $\text{Meet}(X; U_1, \ldots, U_m)$, but $Y_0 \cup Y_1 \cup \cdots \cup Y_m$ must be a linearly independent set. Since $Y_1 \cup \cdots \cup Y_m$ span

$S_1 = U_1 + \cdots + U_m$, it is equivalent to say that Y_0 is an arbitrary independent set in $S_0/S_1 = \mathcal{U}/S_1 = \mathcal{U}/(U_1 + \cdots + U_m)$.

For any such X we have that (3.2.8) holds for any $i \in [m]$ with equality, which proves (3.2.3). Since the $Y_0, Y_1, \ldots, Y_m$ are mutually disjoint (each correspond to elements $x \in X$ with a different meet(x) value), for any i we have $Y_i, \ldots, Y_m$ are mutually disjoint and $Y_i \cup \cdots \cup Y_m$ lie in S_i; therefore (3.2.8) implies that

$$|Y_i \cup \cdots \cup Y_m| = \dim(S_i),$$

and $Y_i \cup \cdots \cup Y_m$ is a basis for S_i; this proves (3.2.4). Since

$$|Y_i| = \dim(S_i) - \dim(S_{i+1}) = \dim(S_i/S_{i+1}),$$

we infer (3.2.5), and combining this with (3.2.7) yields (3.2.6).

Finally, to prove (3.2.7), we have shown $X = Y_0 \cup Y_1 \cup \cdots \cup Y_m$ satisfies (3.2.2) with equality. Thus, X maximizes Meet$(X; U_1, \ldots, U_m)$. Since X is a minimizer for $U_1, \ldots, U_m$ iff X maximizes Meet$(X; U_1, \ldots, U_m)$ over all $X \in \text{Ind}(\mathcal{U})$, we have (3.2.7). □

3.2.3 Decomposing Discoordination into "k-Fold Intersection" Parts

In Section 3.3 we will show that for any subspaces $U_1, U_2, U_3 \subset \mathcal{U}$ of an $\mathbb{F}$-universe, $\mathcal{U}$, the set of all 2-fold intersections

$$U_1 \cap U_2,\ U_1 \cap U_3,\ U_2 \cap U_3$$

are coordinated. The theorems that we will prove will imply a few interesting facts:

1. the images of $U_1 \cap U_2,\ U_1 \cap U_3,\ U_2 \cap U_3$ in $\mathcal{U}/S_3$ (with $S_3 = U_1 \cap U_2 \cap U_3$) are linearly independent,

2. U_1, U_2, U_3 are coordinated iff $[U_1]_{S_2},\ [U_2]_{S_2},\ [U_3]_{S_2}$ are linearly independent (in $\mathcal{U}/S_2$),

3. $$\text{DisCoord}^{\mathcal{U}}(U_1, U_2, U_3) = \text{DisCoord}^{\mathcal{U}/S_2}([U_1]_{S_2}, [U_2]_{S_2}, [U_3]_{S_2}).$$

In this subsection we want to explain that such properties of discoordination hold whenever $U_1, \ldots, U_m$ are subspaces of a universe such that all their k-fold intersections are coordinated. Considering the notation for S_j in Definition 3.2.6, we will show the following two main results: first, for any $j \geq k$, the images in $\mathcal{U}/S_{j+1}$ of the set of all j-fold intersections, meaning sets of the form

$$U_{\{i_1,\ldots,i_j\}} = U_{i_1} \cap \cdots \cap U_{i_j}$$

(with $i_1, \ldots, i_j$ distinct) are linearly independent; this is claim (7) of Theorem 3.2.10 below; second

$$\mathrm{DisCoord}^{\mathcal{U}/S_k}([U_1]_{S_k}, \ldots, [U_m]_{S_k}) \leq \mathrm{DisCoord}^{\mathcal{U}}(U_1, \ldots, U_m)$$

(we don't know if equality generally holds); this is Theorem 3.2.11. As previously mentioned in Subsection 3.1.5, we know that the discoordination of $U_1, \ldots, U_m$ in $\mathcal{U}$ is generally different from that of the their images in $\mathcal{U}/W$ without some conditions on W.

Theorem 3.2.10. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. For $I \subset [m]$ let*

$$U_I = \bigcap_{i \in I} U_i \,.$$

Assume that for some $k \geq 1$, the set of all k-fold intersections,

$$\big\{ U_I \;\big|\; I \subset [m],\ |I| = k \big\},$$

is coordinated by some $Z \in \mathrm{Ind}(\mathcal{U})$. For $j = 0, \ldots, m$ let

$$Z_j = \{ z \in Z \mid \mathrm{meet}(z; U_1, \ldots, U_m) = j \}$$

and

$$Z_{\geq j} = \bigcup_{i \geq j} Z_i = \{ z \in Z \mid \mathrm{meet}(z; U_1, \ldots, U_m) \geq j \}.$$

Then the following statements hold.

1. *For any $I \subset [m]$ with $|I| = j \geq k$, U_I is coordinated by Z.*

2. *For any $I \subset [m]$ with $|I| = j \geq k$, U_I is coordinated by $Z_{\geq j}$,*

$$U_I \cap Z_{\geq j} \text{ is a basis of } U_I.$$

3. *For any $j \geq k$, S_j is coordinated by $Z_{\geq j}$, and $Z_{\geq j}$ is a basis for S_j.*

4. *For any $j \geq k$, Z_j is a basis for S_j relative to S_{j+1}.*

5. *For each $I \subset [m]$ with $|I| = j \geq k$, in $\mathcal{U}/S_{j+1}$, the set $[U_I \cap Z_j]_{S_{j+1}}$ is a basis for $[U_I]_{S_{j+1}}$ of size $|U_I \cap Z_j|$.*

6. *For any $j \geq k$, each element of Z_j is in a unique element of U_I such that $I \subset [m]$ satisfies $|I| = j$, and so Z_j is partitioned into subsets $\{U_I \cap Z_j\}$ with I ranging over all $I \subset [m]$ with $|I| = j$.*

7. *For any $j \geq k$, the images of U_I in $\mathcal{U}/S_{j+1}$ for $I \subset [m]$ such that $|I| = j$, i.e.,*

$$[U_I]_{S_{j+1}} \text{ for } I \subset [m],\ |I| = j$$

   are linearly independent subspaces of S_j/S_{j+1}.

8. *If X is any minimizer of $U_1, \ldots, U_m$, and we set*

$$X_j = \{x \in X \mid \text{meet}(x; U_1, \ldots, U_m) = j\}$$

   and

$$X_{\geq j} = \{x \in X \mid \text{meet}(x; U_1, \ldots, U_m) \geq j\},$$

   then all the above statements hold with Z replaced everywhere by X.

Proof. Most of the implications easily result from the previous ones, often making use of Proposition 3.1.2.

(1): For any $j \geq 2$, consider any $I \subset [m]$ with $|I| = j$: if i_1, i_2 are distinct elements of I and $I_1 = I \setminus \{i_1\}$, $I_2 = I \setminus \{i_2\}$, then

$$U_{I_1} \cap U_{I_2} = \bigcap_{i \in I_1 \cup I_2} U_i \ = \ U_I.$$

Hence each U_I is the intersection of two $U_{I'}$ with $|I'| = |I| - 1$. By assumption Z coordinates all the $U_{I'}$ with $|I'| = k$, so in view of Proposition 3.1.2, Z coordinates all U_I with $|I| = k+1$. Repeated application shows that the same holds with $|I| = k+2, \ldots, m$.

(2): If $|I| = j \geq k$, any element $z \in Z \cap U_I$ meets all U_i with $i \in I$, and hence meet$(z) \geq j$. Thus $Z \cap U_I = Z_{\geq j} \cap U_I$. By (1), Z coordinates U_I, so $Z \cap U_I = Z_{\geq j} \cap U_I$ is a basis for U_I.

(3): S_j is the span of all U_I with $|I| = j$. Since $Z_{\geq j}$ coordinates each such U_I, Proposition 3.1.2 implies that $Z_{\geq j}$ coordinates S_j and that $S_j \cap Z_{\geq j}$ is a basis for S_j. However, each element of $Z_{\geq j}$ meets j of the $U_1, \ldots, U_m$, consequently each element of $Z_{\geq j}$ lies in some U_I with $|I| = j$, and hence also lies in S_j. Therefore, $Z_{\geq j} \cap S_j = Z_{\geq j}$ is a basis for S_j.

(4): By (3), we have $Z_{\geq j+1}, Z_{\geq j}$ are, respectively, bases for S_{j+1}, S_j. It follows that the set $Z_{\geq j} \setminus Z_{\geq j+1}$ is a basis for S_j relative to S_{j+1}. But $Z_{\geq j} \setminus Z_{\geq j+1}$ equals Z_j.

(5): According to (4), in $\mathcal{U}/S_{j+1}$, the vectors in the set $[U_I \cap Z_j]_{S_{j+1}}$ are linearly independent and are of size $|U_I \cap Z_j|$. Since

$$U_I = \text{Span}(Z_{\geq j} \cap U_I),$$

$$[U_I]_{S_{j+1}} = [U_I + S_{j+1}]_{S_{j+1}} = [\text{Span}(Z')]_{S_{j+1}}$$

where

$$Z' = (Z_{\geq j} \cap U_I) \cup Z_{\geq j+1} = (Z_j \cap U_I) \cup Z_{\geq j+1}.$$

Since $Z_{\geq j+1}$ is a basis for S_{j+1} we have that $[Z']_{S_{j+1}} = [Z_j \cap U_I]_{S_{j+1}}$. This implies that $[\text{Span}(Z')]_{S_{j+1}} = [\text{Span}(Z_j \cap U_I)]_{S_{j+1}} = [U_I]_{S_{j+1}}$. Then, $[Z_j \cap U_I]_{S_{j+1}}$ are a set of linearly independent vectors in $\mathcal{U}/S_{j+1}$ and span $[U_I]_{S_{j+1}}$, and hence form a basis for $[U_I]_{S_{j+1}}$.

(6): This is immediate from the definition of Z_j as $z \in Z$ such that meet$(z) = j$.

(7): We have

$$|Z_j| = \sum_{|I|=j} |Z_j \cap U_I|,$$

and so in $\mathcal{U}/S_{j+1}$ we have

$$\dim(S_j/S_{j+1}) = \sum_{|I|=j} \dim^{\mathcal{U}/S_{j+1}}([U_I]_{S_{j+1}}).$$

Since the U_I with $|I| = j$ span all of S_j, the $[U_I]_{S_{j+1}}$ span all of S_j/S_{j+1} in $\mathcal{U}/S_{j+1}$. By (2.5.2) these subspaces are linearly independent.

(8): According to Theorem 3.2.9, for all $j \in [m]$, each X_j is a purely j-th intersection basis, i.e. a basis for S_j relative to S_{j+1} consisting entirely of x with $\text{meet}(x) = j$. It follows that

$$\sum_{|I|=j} |X_j \cap U_I| = |X_j| = \dim(S_j/S_{j+1}).$$

But (3.2.4) implies that $X_{j+1}, \ldots, X_m$ is a basis for S_{j+1}. As a result, $[X_j]_{S_{j+1}}$ is a linearly independent set in $\mathcal{U}/S_{j+1}$ and

$$|X_j \cap U_I| \leq \dim^{\mathcal{U}/S_{j+1}}([U_I]_{S_{j+1}}) = |Z_j \cap U_I|.$$

The Inequality Summation Principle (Proposition 2.2.1) implies that

$$\sum_{|I|=j} |X_j \cap U_I| \leq \sum_{|I|=j} |Z_j \cap U_I|.$$

This is satisfied with equality iff $|X_j \cap U_I| = |Z_j \cap U_I|$ for all I with $|I| = j$. Visibly this inequality is satisfied with equality, since both sides equal $\dim(S_j/S_{j+1})$. Therefore for each I, the image of $X_j \cap U_I$ in $\mathcal{U}/S_{j+1}$ is a basis for $[U_I]_{S_{j+1}}$.

It follows that for all $j \geq k$, each X_j is partitioned into $X_j \cap U_I$ over all $|I|$, such that the image of $X_j \cap U_I$ in $\mathcal{U}/S_{j+1}$ is a basis for $[U_I]_{S_{j+1}}$. This implies (7), from which all of (1)–(6) follow or already have been established. □

Theorem 3.2.11. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. Then for any $k \in [m]$ we have*

$$\text{DisCoord}^{\mathcal{U}/S_k}\big([U_1]_{S_k}, \ldots, [U_m]_{S_k}\big) \leq \text{DisCoord}^{\mathcal{U}}\big(U_1, \ldots, U_m\big). \tag{3.2.10}$$

Furthermore, for any minimizer, X, and $0 \leq j \leq m$ let

$$X_j = \{x \in X \mid \text{meet}(x; U_1, \ldots, U_m) = j\}$$

$$X_{\geq k} = \bigcup_{j \geq k} X_j, \quad X_{<k} = \bigcup_{j<k} X_j.$$

Then (3.2.10) *holds with equality if for some minimizer, X, the following conditions hold:*

1. *$X_{\geq k}$ coordinates $U_i \cap S_k$ for all i,*
2. *$X' = [X_{<k}]_{S_k}$ is a minimizer for $[U_1]_{S_k}, \ldots, [U_m]_{S_k}$; if so, this implies that the same is true for all minimizers, X, of $U_1, \ldots, U_m$, and*
3. *for all $i \in [m]$, $|U_i \cap X_{<k}|$ equals the size of the number of S_k-cosets in $[A_i]_{S_k} \cap [X_{<k}]_{S_k}$ (the first quantity is always bounded above by the second).*

If (3.2.10) *holds with equality, then the above conditions hold for all minimizers.*

Proof. Let X be a minimizer of $U_1, \ldots, U_m$. Then (3.2.4) implies that $X_{\geq k}$ is a basis for S_k, and hence the map $X_{<k}$ to its image, X', in $\mathcal{U}/S_k$ is a bijection, and the X' are linearly independent in $\mathcal{U}/S_k$. Hence for any subspace $W \subset \mathcal{U}$ we have

$$\left|W \cap X_{\geq k}\right| \leq \dim(W \cap S_k),$$

and $X' \in \text{Ind}(\mathcal{U}/S_k)$ is a linearly independent set with

$$|W \cap X_{<k}| \leq \left|[W]_{S_k} \cap X'\right|,$$

where the right most term involving X' counts the size of $[A]_{S_k} \cap X'$ measured in the number of S_k-cosets. Adding the above two displayed inequalities we get

$$|W \cap X| = |W \cap X_{\geq k}| + |W \cap X_{<k}| \leq \dim(W \cap S_k) + \left|[W]_{S_k} \cap X'\right|.$$

Hence

$$\text{DisCoord}^{\mathcal{U}}(U_1, \ldots, U_m) = \sum_{i=1}^{m} \left(\dim(U_i) - |U_i \cap X|\right)$$

$$\geq \sum_{i=1}^{m} \Big(\dim(U_i) - \dim(U_i \cap S_k) - \big|[U_i]_{S_k} \cap X'\big| \Big) \tag{3.2.11}$$

$$= \sum_{i=1}^{m} \Big(\dim^{\mathcal{U}/S_k}\big([U_i]_{S_k}\big) - \big|[U_i]_{S_k} \cap X'\big| \Big) = \mathrm{DisCoord}_{X'}^{\mathcal{U}/S_k}\big([U_1]_{S_k}, \ldots, [U_m]_{S_k}\big)$$

$$\geq \mathrm{DisCoord}^{\mathcal{U}/S_k}\big([U_1]_{S_k}, \ldots, [U_m]_{S_k}\big), \tag{3.2.12}$$

which implies (3.2.10). Furthermore, this inequality is strict unless both (3.2.11) and (3.2.12) hold with equality, which means that conditions (1)–(3) of the theorem statement must hold for (3.2.10) to be satisfied with equality.

If both (3.2.11) and (3.2.12) hold for some minimizer, X, of $U_1, \ldots, U_m$, then we have

$$\mathrm{DisCoord}^{\mathcal{U}}(U_1, \ldots, U_m) = \mathrm{DisCoord}^{\mathcal{U}/S_k}\big([U_1]_{S_k}, \ldots, [U_m]_{S_k}\big),$$

it follows that for *any* minimizer, the two inequalities (3.2.11) and (3.2.12) must also hold with equality, and hence conditions (1)–(3) must hold for all minimizers. □

3.2.4 An Equivalent Discoordination Formula

There is another way to write (3.2.7).

Proposition 3.2.12. *Let $U_1, \ldots, U_m$ be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let notation be as in Definition 3.2.6. Then the discoordination of $U_1, \ldots, U_m$ equals $d_m + d_{m-1} + \cdots + d_1$ where d_j is given by*

$$d_j = \left(\sum_{i=1}^{m} \dim^{\mathcal{U}/S_{j+1}}([U_i \cap S_j]_{S_{j+1}}) \right) - j \dim(S_j/S_{j+1})$$

Proof. For arbitrary

$$0 = S_{m+1} \subset S_m \subset \cdots \subset S_1 \subset S_0 = \mathcal{U},$$

and any subspace $U_i \subset \mathcal{U}$, we have

$$\dim^{\mathcal{U}/S_{j+1}}([U_i \cap S_j]_{S_{j+1}}) = \dim(U_i \cap S_j) - \dim(U_i \cap S_{j+1}).$$

This allows us to write

$$\dim(U_i) = \sum_{j=1}^{m} \dim^{\mathcal{U}/S_{j+1}}([U_i \cap S_j]_{S_{j+1}}).$$

By (3.2.6) and the above equality the proposition follows. □

Since S_j is the span of all j-fold intersections one can intuitively view d_j from the proposition above as a measurement of the "failure" of the images of $U_1 \cap S_j, \ldots, U_m \cap S_j$ in $\mathcal{U}/S_{j+1}$ to be linearly independent in $\mathcal{U}/S_{j+1}$.

3.3 Coordination of Quasi-Increasing Sequences

In this section we prove two main theorems, which we now state. These results will aid in proving Theorem 3.1.4.

Theorem 3.3.1. *Let $\mathcal{U}$ be an $\mathbb{F}$-universe, and let*

$$U_1 \subset \cdots \subset U_s, \quad W_1 \subset \cdots \subset W_t$$

be two sequences of increasing subspaces of $\mathcal{U}$. Then the set of subspaces $U_i \cap W_j$ ranging over all $i \in [s]$ and $j \in [t]$ are coordinated.

A proof of this theorem is given in Subsection 3.3.3. This theorem has the following corollary.

Corollary 3.3.2. *Let $\mathcal{U}$ be an $\mathbb{F}$-universe, and let*

$$U_1 \subset \cdots \subset U_s, \quad W_1 \subset \cdots \subset W_t$$

be two sequences of increasing subspaces of $\mathcal{U}$. Then the subspaces $U_1, \ldots, U_s$, $W_1, \ldots, W_t$ are coordinated.

The corollary is obtained from the theorem by extending the sequences of vector spaces by setting $U_{s+1} = W_{t+1} = \mathcal{U}$; then for all $i \in [s]$, $U_i \cap W_{t+1} = U_i \cap \mathcal{U} = U_i$

and similarly $U_{s+1} \cap W_j = W_j$ for all $j \in [t]$. The corollary is more succinct, since the intersection of any two subspaces coordinated by some basis X is again coordinated by X.

For our analysis of coded caching, we have found the result with $s = 2$ and $t = 1$ helpful. However in simplifying our results we have been able to forgo any use of Theorem 3.3.1.

Note that if $s = t = 2$ and we let $U_2 = W_2 = \mathcal{U}$, Theorem 3.3.1 implies that $U_1 \cap W_1, U_1, W_1$ are coordinated, which is all one needs to prove the dimension formula. Naturally, Theorem 3.3.1 can be viewed as a generalization of the dimension formula.

Theorem 3.3.3. *Let U_1, U_2, U_3 be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. Then the six spaces*

$$U_1 \cap U_2 \cap U_3,\ U_1 \cap U_2,\ U_1 \cap U_3,\ U_2 \cap U_3,\ U_1,\ U_2$$

are coordinated.

This theorem will be crucial in showing Theorem 3.1.4 and a proof is given in Subsection 3.3.4. It turns out that both theorems can be proven by a strategy that generalizes the proof of the dimension formula.

3.3.1 Quasi-Increasing Sequences

If $U_1 \subset \cdots \subset U_m$ are a set of increasing subspaces of some universe, then this sequence is coordinated: consider a basis for U_1, and successively extend that basis to a basis for U_2 and so on. In this subsection we give a more general situation where a similar strategy works.

Definition 3.3.4. Let $U_1, \ldots, U_m$ be a sequence of vector spaces in some universe. For $r = 2, 3, \ldots, m$, we say that this sequence is *quasi-increasing in position r* if whenever

$$v_r = v_1 + \cdots + v_{r-1} \text{ for } v_1 \in U_1, \ldots, v_r \in U_r, \tag{3.3.1}$$

one also has

$$v_r = v_1' + \cdots + v_{r-1}' \text{ for some } v_1' \in U_1, \ldots, v_{r-1}' \in U_{r-1},$$

such that

$$v_i' \neq 0 \;\Rightarrow\; U_i \subset U_r$$

(i.e., if $i < r$ and $U_i \not\subset U_r$ then $v_i' = 0$). If this condition holds for all $r = 2, \ldots, m$, we say that $U_1, \ldots, U_m$ is *quasi-increasing*.

To be quasi-increasing in position r is equivalent to

$$\big(U_r \cap (U_1 + \cdots + U_{r-1})\big) \subset \sum_{i<r \text{ and } U_i \subset U_r} U_i\,; \tag{3.3.2}$$

the reverse inclusion is clear, so we can replace $\subset$ with $=$ if we need to. Furthermore, any increasing sequence $U_1 \subset \cdots \subset U_m$ is also quasi-increasing, since whenever $i < r$ we have $U_i \subset U_r$.

Example 3.3.5. Let A, B be any vector spaces in some universe, and let $U_1 = A \cap B$, $U_2 = A$, and $U_3 = B$. Then $U_1 \subset U_2$ and $U_1 \subset U_3$, but $U_2 \not\subset U_3$. However, if $v_1 \in U_1$, $v_2 \in U_2$, and $v_3 \in U_3$ with

$$v_3 = v_1 + v_2,$$

then in fact v_3 lies in U_1 (one sees this by first noting that $v_2 = v_3 - v_1$, and since $v_1, v_3 \in B$ then also $v_2 \in B$; since $v_2 \in A$ then $v_2 \in A \cap B = U_1$). Hence the above sequence is quasi-increasing, but not generally increasing.

Example 3.3.5 is the essential step in proving the dimension formula: namely we let X_1 be a basis for $U_1 = A \cap B$, X_2 a minimal set such that $X_1 \cup X_2$ spans A, and X_3 a minimal set such that $X_1 \cup X_3$ spans B. We then see that $X_1 \cup X_2$ is a basis for $U_2 = A$; to show that $X_1 \cup X_2 \cup X_3$ is a basis, we need to show that there is no nontrivial relation between the vectors of X_3 and those of X_1 and X_2; if there is such a relation we would have

$$v_1 + v_2 = v_3 \neq 0$$

where each v_i is a linear combination of vectors in X_i, but then (as shown in Example 3.3.5) we will have $v_3 \in U_1$, which contradicts the fact that $X_1 \cup X_3$ are linearly independent.

The theorem below strengthens the method used to prove the dimension formula.

Theorem 3.3.6. *Any quasi-increasing sequence is coordinated. In more detail, let $U_1, \ldots, U_m$ be a sequence of quasi-increasing subspaces in some universe. Let X_1 be any basis for U_1, and inductively on $i = 2, \ldots, m$, let X_i be a minimal size set of vectors that when added to*

$$X' = \bigcup_{j \text{ s.t. } U_j \subset U_i} X_j,$$

$X_i \cup X'$ spans U_i. Then $X_1, \ldots, X_m$ are mutually disjoint and $X = X_1 \cup \cdots \cup X_m$ coordinate $U_1, \ldots, U_m$, and, more specifically, for each i we have

$$X \cap U_i = \bigcup_{j \text{ s.t. } U_j \subseteq U_i} X_j$$

is a basis for U_i.

Proof. We prove this by induction on m. The base case $m = 1$ is clear since X_1 is simply a basis for U_1.

Now say that the theorem holds for some value of $m \geq 1$, let $U_1, \ldots, U_m, U_{m+1}$ be a quasi-increasing sequence, and $X_1, \ldots, X_m$ be sets of vectors as in the theorem. Let $I = \{i \in [m] \mid U_i \subset U_{m+1}\}$. By Proposition 3.1.2, $X' = \bigcup_{i \in I} X_i$ coordinates

$$U' = \sum_{i \in I} U_i; \tag{3.3.3}$$

since $U_i \subset U_{m+1}$ for all $i \in I$, we have $U' \subset U_{m+1}$. Let X_{m+1} be as specified in the theorem. Then the vectors $X' \cup X_{m+1}$ are linearly independent (and X_{m+1} is disjoint from X'). By assumption, $X_1 \cup \cdots \cup X_m$ are (mutually disjoint and) linearly independent. Then if $X_1 \cup \cdots \cup X_{m+1}$ is not a linearly independent set of vectors (or if X_{m+1} is not distinct from $X_1, \ldots, X_m$), we have

$$v_{m+1} = v_1 + \cdots + v_m,$$

where each v_i is in the span of X_i and v_{m+1} is nonzero. But then we have

$$v_{m+1} = v'_1 + \cdots + v'_m,$$

where $v_i' \neq 0$ only if $i \in I$; hence $v_1' + \cdots + v_m' \in U'$, which contradicts the fact that $X' \cup X_{m+1}$ are linearly independent. $\square$

The main subtlety in the proof of Theorem 3.3.6 is considering U' in (3.3.3) which is the sum of all subspaces that lie in U_{m+1}. The dimension formula works with U_1, U_2, U_3 as in Example 3.3.5, where U_1 is the sole subspace that lies in U_3.

We remark that not all sets of coordinated subspaces in some universe form a quasi-increasing sequence. Consider, in $\mathbb{F}^3$ where e_1, e_2, e_3 denote the standard basis vectors, the subspaces

$$U_1 = \text{Span}(e_1),\ U_2 = \text{Span}(e_1, e_2),\ U_3 = \text{Span}(e_2, e_3);$$

these subspaces are coordinated by $X = \{e_1, e_2, e_3\}$, but there is no ordering of them that results in a quasi-increasing sequence. For instance with the sequence U_1, U_2, U_3, take $v_3 = e_2 \in U_3$, then $v_3 = v_2 \in U_2$ where $v_2 \neq 0$ but $U_2 \not\subset U_3$. The other orderings follow a similar argument.

3.3.2 Maximal Index Sets

Say that we are trying to prove that a given sequence $U_1, \ldots, U_m$ of subspaces in a universe is quasi-increasing. Meaning, given any r between 2 and m and any equation

$$v_r = v_{r-1} + \cdots + v_1,$$

we wish to find $v_1', \ldots, v_{r-1}'$ as in Definition 3.3.4, i.e., whose sum is also v_r, with $v_i' \in U_i$ for $i \in [r-1]$, but such that $v_i' = 0$ if $U_i \not\subset U_r$. In practice one can simplify this task by noting that if $i < j < r$ and $U_i \subset U_j$, then we can always assume that $v_i = 0$, by replacing v_j with $v_j + v_i$. This idea leads to the following simplification.

Definition 3.3.7. Let $U_1, \ldots, U_m$ be a sequence of vector spaces in a universe, $\mathcal{U}$. For $r = [m]$ we define the *r-maximal index set*, denoted M_r, to be

$$M_r = [r] \setminus \{i \in [r] \mid U_i \subset U_j \text{ for some } j \neq i\} \tag{3.3.4}$$

(equivalently, $i \in M_r$ if U_i is a maximal subspace under inclusion among $U_1, \ldots, U_r$).

Proposition 3.3.8. *Let $U_1, \ldots, U_r$ be a sequence of vector spaces in a universe, $\mathcal{U}$. Then*

$$U_1 + \cdots + U_r = \sum_{i \in M_r} U_i, \tag{3.3.5}$$

i.e., any sum of elements in $U_1, \ldots, U_r$ can be written as a sum of those U_i that are maximal under inclusion.

Proof. In case $r = 1$, we have $M_r = M_1 = \{1\}$. Now we argue by induction on r: say the proposition holds for r replaced by $r-1$ for some $r \geq 2$; hence

$$U_1 + \cdots + U_{r-1} = \sum_{i \in M_{r-1}} U_i.$$

If $r \in M_r$, then $M_r = \{r\} \cup (M_{r-1} \setminus I')$ where I' is the set of $i < r$ with $U_i \subset U_r$. Then

$$\sum_{i \in I'} U_i \subset U_r,$$

so we have

$$U_1 + \cdots + U_{r-1} \ \subset \ U_r + \sum_{i \in (M_{r-1} \setminus I')} U_i \ = \ \sum_{i \in M_r} U_i.$$

Adding U_r to both sides we obtain

$$U_1 + \cdots + U_{r-1} + U_r \ \subset \ \sum_{i \in M_r} U_i.$$

But since $M_r \subset [r]$, the reverse inclusion is clear, and hence we obtain (3.3.5). Otherwise $r \notin M_r$, as a result $M_{r-1} = M_r$ and $U_r \subset U_k$ for some $k \in [r-1]$. Then

$$U_1 + \cdots + U_{r-1} + U_r \ = \ U_1 + \cdots + U_{r-1} + U_k \ = \ U_1 + \cdots + U_{r-1} \ = \ \sum_{i \in M_{r-1}} U_i,$$

and since $M_{r-1} = M_r$ we again conclude (3.3.5). □

3.3.3 Coordination of Two Sequences of Increasing Subspaces

In this subsection we prove Theorem 3.3.1. According to Theorem 3.3.6, it suffices to prove the following stronger theorem.

Theorem 3.3.9. *Let $\mathcal{U}$ be an $\mathbb{F}$-universe, and let*

$$A_1 \subset \cdots \subset A_s, \quad B_1 \subset \cdots \subset B_t$$

be two sequences of increasing subspaces of $\mathcal{U}$. Let us order the $s \cdot t$ subspaces of the form $A_i \cap B_j$ as follows:

$$U_1 = A_1 \cap B_1, \ldots, U_s = A_s \cap B_1, U_{s+1} = A_1 \cap B_2, \ldots, U_{2s} = A_s \cap B_2, \ldots, U_{st} = A_s \cap B_t$$

(i.e., for all $i \in [s]$ and $j \in [t]$, we set $U_{i+s(j-1)} = A_i \cap B_j$). Then $U_1, \ldots, U_{st}$ is a quasi-increasing sequence of subspaces.

Proof. Let us prove the theorem by induction on t. For $t = 1$, the sequence $U_1, \ldots, U_s$ is increasing, and therefore quasi-increasing.

Now say that we know the theorem holds for $t-1$ where $t \geq 2$. For any $i \in [s]$, let $r = i + s(t-1)$, where $U_r = A_i \cap B_t$, and let us verify the condition in Definition 3.3.4. Note that

$$U_j \subset A_s \cap B_{t-1} = U_{s(t-1)} \quad \text{for } j \in [s(t-1)],$$

and that

$$U_{1+s(t-1)} = A_1 \cap B_t \subset U_{2+s(t-1)} = A_2 \cap B_t \subset \cdots \subset U_{st} = A_s \cap B_t.$$

Then for M_{r-1}, the $(r-1)$-maximal index set from Definition 3.3.7, and $i \in [s]$ we have

$$M_{r-1} = \big\{ s(t-1),\ s(t-1) + (i-1) \big\}.$$

Consider the case $i \geq 2$ (where $|M_{r-1}| = 2$). For $v_i \in U_i$, if $v_r = v_1 + \cdots + v_{r-1}$, then by Proposition 3.3.8 we also have

$$v_r = w_1 + w_2$$

with $w_1 \in A_s \cap B_{t-1}$ and $w_2 \in A_{i-1} \cap B_t$. But then $w_1 = v_r - w_2$ is in A_i, since $v_r, w_2 \in A_i \cap B_t$; hence $w_1 \in A_i$, and therefore $w_1 \in A_i \cap B_{t-1}$. But both $A_{i-1} \cap B_t$ and $A_1 \cap B_{t-1}$ are subsets of $A_i \cap B_t = U_r$ that occur in the list $U_1, \ldots, U_{r-1}$. This

establishes the quasi-increasing condition in $U_1, \ldots, U_r$ for the values of r with $i \geq 2$.

In case $i = 1$ we have that $M_{r-1} = \{s(t-1)\}$ and the same argument works (without a w_2). □

3.3.4 The Coordinated Parts of Three Subspaces

In this subsection we will prove Theorem 3.3.3 and give a related lemma which will be useful for proving Theorem 3.1.8. Again, to prove Theorem 3.3.3, it will suffice to prove this stronger result.

Theorem 3.3.10. *Let A, B, C be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$. Then the sequence*

$$U_1 = A \cap B \cap C,\ U_2 = A \cap B,\ U_3 = A \cap C,\ U_4 = B \cap C,\ U_5 = A,\ U_6 = B$$

is quasi-increasing.

Proof. We need to verify that the sequence is quasi-increasing in positions $r = 2, 3, \ldots, 6$. Let $v_i \in U_i$ for $i \in [6]$.

For $r = 2$ we have $U_1 \subset U_2$ so the condition holds.

For $r = 3$, $U_2 \cap U_3 = U_1$, so the verification is the same as for the dimension formula.

For $r \geq 4$, since $U_1 \subset U_2$, we can omit v_1 from all equations (3.3.1).

For $r = 4$, $U_4 = B \cap C$, consider an equation $v_4 = v_2 + v_3$. Then since $v_4, v_2 \in B$, the equation $v_3 = v_4 - v_2 \in B$ shows that $v_3 \in U_3 \cap B = U_1$. Similarly one shows $v_2 \in U_1$. Thus, we may take $v_1' = v_2 + v_3 \in U_1$ and have $v_4 = v_1' \in U_1 \subset U_4$.

For $r = 5$, we consider an equation

$$v_5 = v_2 + v_3 + v_4.$$

Since $v_2, v_3, v_5 \in A$, then $v_4 \in A$ and hence $v_4 \in A \cap U_4 = U_1 \subset U_5$. Furthermore $U_1, U_2, U_3 \subset A = U_5$, therefore the verification is complete.

For $r = 6$, since $U_1, U_2, U_3 \subset A = U_5$, it suffices to consider the equation

$$v_6 = v_4 + v_5.$$

Since $v_4, v_6 \in B$ we have $v_5 \in B$ and hence $v_5 \in B \cap A = U_2 \subset B = U_6$. Given that $U_2, U_4 \subset U_6$, the verification is complete. □

The same idea can be used to show that for any $U_1, \ldots, U_m$, the set of all $(m-1)$-fold intersections are coordinated. However, for $m \geq 4$, the set of all $(m-2)$-fold intersections can be discoordinated. For example, in $\mathbb{F}^3$ let

$$U_1 = \text{Span}(e_1, e_2),\ U_2 = \text{Span}(e_1, e_3),\ U_3 = \text{Span}(e_2, e_1 + e_3),\ U_4 = \text{Span}(e_3),$$

where the 2-fold intersections include the one dimensional spaces spanned by e_1, e_2, e_3, and $e_1 + e_3$. If $m \geq 5$ we can set $U_i = \mathcal{U}$ for $i \geq 5$ and have a similar problem with the $(m-2)$-fold intersections.

The following lemma is helpful in showing Theorem 3.1.8 and is used in Section 3.5.

Lemma 3.3.11. *Let A, B, C, D be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$ such that $D \subset (A \cap B)$. Then the sequence*

$$U_1 = A \cap B \cap C \cap D = C \cap D,\ U_2 = A \cap B \cap C,\ U_3 = D,$$

$$U_4 = A \cap B,\ U_5 = A \cap C,\ U_6 = B \cap C,\ U_7 = A,\ U_8 = B.$$

is quasi-increasing.

Proof. We need to verify that the sequence is quasi-increasing in positions $r = 2, 3, \ldots, 8$. Let $v_i \in U_i$ for $i \in [8]$.

For $r = 2$, since $U_1 \subset U_2$, the sequence is quasi-increasing up to $r = 2$.

For $r = 3$, $U_2 \cap U_3 = U_1$, so the verification is the same as for the dimension formula.

For $r = 4$, $U_i \subset U_4$ for $i \in [3]$, then the sequence is quasi-increasing up to $r = 4$.

For $r \geq 5$, since $U_i \subset U_4$ for $i \in [3]$, we can omit v_1, v_2, v_3 from all equations (3.3.1), thus using the same argument as the proof of Theorem 3.3.10 we have that the sequence is quasi-increasing. □

3.3.5 Strongly Quasi-Increasing Sequences

We remark that in Theorem 3.3.9 and Theorem 3.3.10 the sequences $U_1, \ldots, U_m$ have a stronger property: namely if for each r we set

$$M'_r = \{i \in [m] \mid i \neq r \quad \text{and} \quad U_r \not\subset U_i\}$$

(so M'_r can contain i greater than r), then we have

$$U_r \cap \left(\sum_{i \in M'_r} U_i \right) \subset \sum_{i \in M'_r} U_i.$$

In other words, if we rearrange the $U_1, \ldots, U_m$ in any order $U'_1, \ldots, U'_m$ such that $i < j$ implies $U_j \not\subset U_i$ (i.e., $U'_1, \ldots, U'_m$ is any total ordering that extends the partial ordering on $U_1, \ldots, U_m$ under inclusion), then we still have that $U'_1, \ldots, U'_m$ is quasi-increasing. In this case we say $U_1, \ldots, U_m$ is *strongly quasi-increasing*. This is a property of $U_1, \ldots, U_m$ viewed as a poset under inclusion.

The quasi-increasing sequence used in Theorem 3.3.9 is strongly quasi-increasing. The argument is similar for the sequence in Theorem 3.3.10, therefore a detailed argument is omitted.

Consider the sets $A_i \cap B_j$ in Theorem 3.3.9: if $A_i \cap B_j \not\subset A_{i'} \cap B_{j'}$ (and the $A_1, \ldots, A_m$ are distinct, as well as the $B_1, \ldots, B_t$), then either $i' < i$ or $j' < j$; hence $A_{i'} \cap B_{j'}$ is a subset of either $A_{i-1} \cap B_t$ (and $i \geq 2$) or a subset of $A_s \cap B_{j-1}$ (and $j \geq 2$). But if

$$v_r = w_1 + w_2$$

with $v_r \in A_i \cap B_j$, $w_1 \in A_{i-1} \cap B_t$, and $w_2 \in A_s \cap B_{j-1}$, then writing $w_2 = v_r - w_1$ shows that $w_2 \in A_i$, and hence $w_2 \in A_i \cap B_{j-1}$ which lies in $A_i \cap B_j$; similarly one can show w_1 lies in B_j, which implies $w_1 \in A_{i-1} \cap B_j \subset A_i \cap B_j$.

We are not certain whether this strong quasi-increasing property is an accident in these instances or holds whenever a sequence is quasi-increasing.

3.4 The Main Discoordination Theorem

Building upon what we have done in the previous sections, the goal of this section is to prove Theorem 3.1.4.

3.4.1 The $S_2 = 0$ Case

Consider the case where the S_2 (from Definition 3.2.6) of three subspaces of a universe is zero, i.e., when the triple and double intersections of the subspaces are empty. We will show the following theorem under this condition.

Theorem 3.4.1. *Let A, B, C be any subspaces of an $\mathbb{F}$-universe $\mathcal{U}$ such that $A \cap B = A \cap C = B \cap C = 0$. Let $m = \text{DisCoord}(A, B, C)$. Then:*

1. $\dim(A+B) = \dim(A) + \dim(B)$*;*
2. $m = \dim\big((A+B) \cap C\big)$*;*
3. *there are bases $a_1, \ldots, a_{m_1}$ of A, $b_1, \ldots, b_{m_2}$ of B, and $c_1, \ldots, c_{m_3}$ of C, such that $m_1, m_2, m_3 \geq m$, and we have*

$$c_i = a_i + b_i \quad \text{for} \quad i \in [m],$$

which implies

$$a_1, \ldots, a_{m_1},\ b_1, \ldots, b_{m_2},\ c_{m+1}, \ldots, c_{m_3} \tag{3.4.1}$$

is a basis for $A + B + C$.

Proof. By the dimension formula, and since $A \cap B = 0$, we have

$$\dim(A+B) = \dim(A) + \dim(B).$$

By assumption we have $S_3 = S_2 = 0$ and since $S_1 = A + B + C$, by Theorem 3.1.9 we have

$$\text{DisCoord}(A, B, C) = \dim(A) + \dim(B) + \dim(C) - \dim(A+B+C) = m, \tag{3.4.2}$$

and since $\dim(A)+\dim(B)=\dim(A+B)$,

$$m=\dim(A+B)+\dim(C)-\dim(A+B+C). \tag{3.4.3}$$

The dimension formula applied to $A+B$ and C gives us

$$\dim(A+B)+\dim(C)=\dim\big((A+B)\cap C\big)+\dim(A+B+C),$$

which combined with (3.4.3) implies $m=\dim\big((A+B)\cap C\big)$.

Let $c_1,\ldots,c_m$ be a basis for $(A+B)\cap C$; since each c_i also lies in $A+B$, we may write each c_i as a_i+b_i. We claim that $a_1,\ldots,a_m$ are linearly independent, for if not then for some $\gamma_1,\ldots,\gamma_m\in\mathbb{F}$ we have

$$\gamma_1 a_1+\cdots+\gamma_m a_m=0,$$

and hence

$$\gamma_1 c_1+\cdots+\gamma_m c_m=-\gamma_1 b_1-\cdots-\gamma_m b_m;$$

but this is impossible, since the left-hand-side is a nonzero element of C, and the right-hand-side is an element of B, which would imply that $C\cap B$ contains a nonzero element, contrary to the assumption of the theorem. Similarly the $b_1,\ldots,b_m$ are linearly independent.

By basis extension, we may extend $a_1,\ldots,a_m$ to get a basis, $a_1,\ldots,a_{m_1}$ of A with $m_1\geq m$. Similarly we extend the $b_1,\ldots,b_m$ to get a basis $b_1,\ldots,b_{m_2}$ of B, with $m_2\geq m$. Since $(A+B)\cap C$ is a subspace of dimension m in C, with a basis $c_1,\ldots,c_m$, we may extend this to get a basis $c_1,\ldots,c_{m_3}$ of C with $m\geq m_3$. From (3.4.2) and since by construction m_1,m_2,m_3 are the dimensions of A,B,C we have

$$\dim(A+B+C)=\dim(A)+\dim(B)+\dim(C)-m=m_1+m_2+m_3-m.$$

Since $A+B+C$ is spanned by $a_1,\ldots,a_{m_1},\ b_1,\ldots,b_{m_2},\ c_{m_3-m+1},\ldots,c_{m_3}$ and these are $m_1+m_2+m_3-m$ linearly independent vectors, we have that the collection of vectors in (3.4.1) is a basis for $A+B+C$. □

3.4.2 The Lifting Lemma

Before we prove Theorem 3.1.4, it is helpful to extract a simple ingredient of the proof that is conceptually important.

Lemma 3.4.2 (The Lifting Lemma)**.** *Let A, B, C be subspaces of an $\mathbb{F}$-universe, $\mathcal{U}$, and let*

$$S_2 = \mathcal{S}_2(A,B,C) = A \cap B + A \cap C + B \cap C.$$

If for some $\tilde{a}, \tilde{b}, \tilde{c}$ we have

$$[\tilde{a} + \tilde{b}]_{S_2} = [\tilde{c}]_{S_2},$$

then there exist $a \in A$, $b \in B$, and $c \in C$ such that

$$a + b = c$$

and

$$[a]_{S_2} = [\tilde{a}]_{S_2},\ [b]_{S_2} = [\tilde{b}]_{S_2},\ [c]_{S_2} = [\tilde{c}]_{S_2}. \tag{3.4.4}$$

In particular we have

$$[A+B]_{S_2} \cap [C]_{S_2} = \big[(A+B) \cap C\big]_{S_2}.$$

Proof. Suppose $[\tilde{a} + \tilde{b}]_{S_2} = [\tilde{c}]_{S_2}$, then

$$[\tilde{a} + \tilde{b} - \tilde{c}]_{S_2} = [0]_{S_2}$$

and therefore

$$\tilde{a} + \tilde{b} - \tilde{c} = v_1 + v_2 + v_3$$

for some $v_1 \in A \cap B$, $v_2 \in A \cap C$, and $v_3 \in B \cap C$. Then $v_2, v_3 \in C$ and as a result

$$c = \tilde{c} + v_2 + v_3 \in C.$$

Similarly $v_1 \in A$, thus $a = \tilde{a} - v_1 \in A$. Taking $b = \tilde{b}$ we get $a + b = c$. Since each $v_1, v_2, v_3 \in S_2$, we have (3.4.4).

It is immediate that

$$\big[(A+B)\cap C\big]_{S_2} \subset [A+B]_{S_2}\cap[C]_{S_2};$$

to prove the reverse inclusion we note that an element of the right-hand-side of the above equation is a class $[\tilde{c}]_{S_2}$ with $\tilde{c}\in C$ which is also a class of the form $[\tilde{a}+\tilde{b}]_{S_2}$; but then there are $a\in A$, $b\in B$, and $c\in C$ with $c=a+b$ as in the previous paragraph, and hence $c\in C\cap(A+B)$ such that $[c]_{S_2}=[\tilde{c}]_{S_2}$. □

3.4.3 Proof of the Main Theorem Regarding Three Subspaces

Proof of Theorem 3.1.4. For ease of notation let us use A,B,C instead of U_1,U_2,U_3 to denote the three subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$. According to Theorem 3.3.3, the subspaces

$$A\cap B\cap C,\ A\cap B,\ A\cap C,\ C\cap B$$

are coordinated; so let X be a basis for

$$S_2=(A\cap B)+(A\cap C)+(B\cap C)$$

that coordinates these subspaces. Consider in $\mathcal{U}/S_2$ the vector subspaces $A'=[A]_{S_2}$, $B'=[B]_{S_2}$, $C'=[C]_{S_2}$; apply Theorem 3.4.1 to these three subspaces (whose two-fold intersections vanish). Let

$$a'_1,\ldots,a'_{m_1},\ b'_1,\ldots,b'_{m_2},\ c'_1,\ldots,c'_{m_3}$$

be the respective bases for A',B',C' with $c'_i=a'_i+b'_i$ for $i\in[m]$, where $m=\mathrm{DisCoord}^{\mathcal{U}/S_2}(A',B',C')$.

Each a'_j is an S_2-coset, so for each $j\in[m_1]$ pick an arbitrary $\tilde{a}_j\in\mathcal{U}$ with $[\tilde{a}_j]_{S_2}=a'_j$, similarly pick $\tilde{b}_k$ and $\tilde{c}_l$ with $k\in[m_2]$ and $l\in[m_3]$. By Lemma 3.4.2, for each $i\in[m]$ there exist a_i,b_i,c_i whose S_2-coset is the same as $\tilde{a}_i,\tilde{b}_i,\tilde{c}_i$ respectively, and satisfy $a_i+b_i=c_i$.

For $j > m$, let $a_j = \tilde{a}_j$, and similarly define b_k and c_l for $k, l > m$. Setting

$$X' = \{a_1, \ldots, a_{m_1},\ b_1, \ldots, b_{m_2},\ c_{m+1}, \ldots, c_{m_3}\}$$

we see that $X \cup X'$ is a basis for $\mathcal{U}$ since X' is a basis of $\mathcal{U}$ relative to S_2 and X is a basis of S_2. Let

$$X_2 = \{a_1, \ldots, a_m, b_1, \ldots, b_m\}$$

and

$$X_1 = \ (X \cup X') \setminus X_2 \ = \ X \cup (X' \setminus X_2).$$

Set $\mathcal{U}_1 = \mathrm{Span}(X_1)$ and $\mathcal{U}_2 = \mathrm{Span}(X_2)$. Then X_1, X_2 are disjoint sets whose union is a basis of $\mathcal{U}$, and hence $\mathcal{U}_1$,$\mathcal{U}_2$ form a decomposition of $\mathcal{U}$.

We have

$$\dim(A \cap \mathcal{U}_1) = \dim(A \cap S_2) + \dim^{\mathcal{U}/S_2}(A) - m, \quad \dim(A \cap \mathcal{U}_2) = m,$$

and the same equalities with B and C replacing A. Then A, B, C factor through this decomposition (note that X_2 contains $c_i = a_i + b_i$ for $i \in [m]$).

Since $\dim(\mathcal{U}_2) = 2m$, and if $\mu : \mathcal{U}_2 \to \mathbb{F}^2 \otimes \mathbb{F}^m$ is the isomorphism taking a_i to $e_1 \otimes e_i$ and b_i to $e_2 \otimes e_i$, then μ takes c_i to $(e_1 + e_2) \otimes e_i$ for all $i \in [m]$. Hence μ satisfies the required condition of Theorem 3.1.4.

Finally, we claim that X_1 coordinates $A \cap \mathcal{U}_1, B \cap \mathcal{U}_1, B \cap \mathcal{U}_1$. To check this for $A \cap \mathcal{U}_1$, we note that $A \cap (X' \setminus X_2)$ is of size at least $m_1 - m$, and thus

$$|A \cap X_1| \ = \ |A \cap X| + |A \cap (X' \setminus X_2)| \ \geq \ \dim(A \cap S_2) + m_1 - m.$$

Since $m_1 = \dim^{\mathcal{U}/S_2}([A]) = \dim(A) - \dim(A \cap S_2)$, we have

$$|A \cap X_1| \geq \dim(A) - m.$$

Given that A factors through the decomposition, we know

$$|A \cap X_1| \ \leq \ \dim(A \cap \mathcal{U}_1) \ = \ \dim(A) - \dim(A \cap \mathcal{U}_2).$$

Since $\dim(A \cap \mathcal{U}_2) = m$, we have $|A \cap X_1| \leq \dim(A) - m$, as a consequence

$$|A \cap X_1| \;=\; \dim(A) - m \;=\; \dim(A \cap \mathcal{U}_1).$$

This shows that X_1 coordinates $A \cap \mathcal{U}_1$. Similarly we have $B \cap (X' \setminus X_2)$ and $C \cap (X' \setminus X_2)$ are at least of size $m_2 - m$ and $m_3 - m$ respectively, with a similar argument we conclude that

$$|B \cap X_1| = \dim(B) - m = \dim(B \cap \mathcal{U}_1), \quad |C \cap X_1| = \dim(C) - m = \dim(C \cap \mathcal{U}_1).$$

Therefore X_1 also coordinates $B \cap \mathcal{U}_1$ and $C \cap \mathcal{U}_1$. □

3.5 Proof of Theorem 3.1.8

The following lemma is useful in showing Theorem 3.1.8.

Lemma 3.5.1. *Let A, B, C be coordinated subspaces of a universe, $\mathcal{U}$, and let $D \subset A \cap B$. Then A, B, C, D are coordinated in $\mathcal{U}$ and hence $[A]_D, [B]_D, [C]_D$ are coordinated in $\mathcal{U}/D$.*

Proof. We will first show that A, B, C, D are coordinated, then use a coordinating basis for A, B, C, D to construct a coordinating basis for $[A]_D, [B]_D, [C]_D$.

Let $S_2 = (A \cap B) + (A \cap C) + (B \cap C)$, by Theorem 3.3.3 all triple and double intersections of A, B, C are coordinated.

Let X be a basis for $\mathcal{U}$ which coordinates A, B, C; by Theorem 3.1.4, X coordinates S_2. Let $X_1 = X \cap S_2$ and $X_2 = X \setminus X_1$. We can decompose $\mathcal{U}$ into $\mathcal{U}_1 = \text{Span}(X_1) = S_2$ and $\mathcal{U}_2 = \text{Span}(X_2)$. A, B, C, D factor through this decomposition. Let $A_1 = A \cap \mathcal{U}_1$ and $A_2 = A \cap \mathcal{U}_2$, similarly define B_i, C_i. Note that X_2 is a coordinating basis for A_2, B_2, C_2. Since $D \subset (A \cap B) \subset S_2 = \mathcal{U}_1$, $D = D \cap \mathcal{U}_1$ and $D \cap \mathcal{U}_2 = 0$.

By Lemma 3.3.11 and Theorem 3.3.6 we know

$$D \cap C,\; A \cap B \cap C,\; D,\; A \cap B,\; A \cap C,\; B \cap C$$

are coordinated; let X_1' be a coordinating basis for these subspaces, note that X_1' is

basis for S_2 and coordinates A_1, B_1, C_1, and D. Then $X' = X_1' \cup X_2$ is a coordinating basis for A, B, C, D.

Let $\hat{X} = X' \setminus (X' \cap D) = \big(X_1' \setminus (X_1' \cap D)\big) \cup X_2$, then $\hat{X}$ is basis for $\mathcal{U}$ relative to D. Furthermore,

$$|\hat{X} \cap A| = |X' \cap A| - |X' \cap D \cap A| = \dim(A) - \dim(A \cap D) = \dim^{\mathcal{U}/D}([A]_D).$$

A similar statement holds for $\hat{X} \cap B$ and $\hat{X} \cap C$, therefore the image of $\hat{X}$ in $\mathcal{U}/D$ is a coordinating basis for $[A]_D, [B]_D$, and $[C]_D$. □

Proof of Theorem 3.1.8. For ease of notation let us use A, B, C, D instead of U_1, U_2, U_3, W to denote the four subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$. Theorem 3.1.4 asserts that $\mathcal{U}$ has a decomposition $\mathcal{U}_1, \mathcal{U}_2$ through which A, B, C factor, such that

1. $A_1 = A \cap \mathcal{U}_1$, $B_1 = B \cap \mathcal{U}_1$, and $C_1 = C \cap \mathcal{U}_1$ are coordinated subspaces of $\mathcal{U}_1$;
2. there is an isomorphism $\mathcal{U}_2 \to \mathbb{F}^2 \otimes \mathbb{F}^m$ that takes $A_2 = A \cap \mathcal{U}_2$, $B_2 = B \cap \mathcal{U}_2$, and $C_2 = C \cap \mathcal{U}_2$, respectively, to

$$\mathrm{Span}(e_1) \otimes \mathbb{F}^m,\ \mathrm{Span}(e_2) \otimes \mathbb{F}^m,\ \mathrm{Span}(e_1 + e_2) \otimes \mathbb{F}^m,$$

   where $\mathrm{DisCoord}^{\mathcal{U}_2}(A_2, B_2, C_2) = \mathrm{DisCoord}^{\mathcal{U}}(A, B, C)$.

Hence $\dim(A_2 \cap B_2) = \dim\Big(\big(\mathrm{Span}(e_1) \otimes \mathbb{F}^m\big) \cap \big(\mathrm{Span}(e_2) \otimes \mathbb{F}^m\big)\Big) = 0$, which implies $A_2 \cap B_2 = 0$. As a result we have

$$A \cap B = (A_1 \cap B_1) + (A_2 \cap B_2) = A_1 \cap B_1.$$

Thus, $D \subset A \cap B$ means that $D \subset (A_1 \cap B_1) \subset \mathcal{U}_1$. In this case $\mathcal{U}/D$ decomposes as $\mathcal{U}_1/D$, $\mathcal{U}_2$ and by Theorem 3.1.10 we have that

$$\begin{aligned}\mathrm{DisCoord}^{\mathcal{U}/D}([A]_D, [B]_D, [C]_D) =& \mathrm{DisCoord}^{\mathcal{U}_1/D}([A_1]_D, [B_1]_D, [C_1]_D) \\ &+ \mathrm{DisCoord}^{\mathcal{U}_2}(A_2, B_2, C_2) \\ =& \mathrm{DisCoord}^{\mathcal{U}_1/D}([A_1]_D, [B_1]_D, [C_1]_D) \\ &+ \mathrm{DisCoord}^{\mathcal{U}}(A, B, C).\end{aligned}$$

Since A_1, B_1, C_1 are coordinated subspaces of $\mathcal{U}_1$ and D is some linear subspace of $\mathcal{U}_1$ such that $D \subset (A_1 \cap B_1)$ by Lemma 3.5.1 we have that $[A_1]_D, [B_1]_D, [C_1]_D$ are coordinated subspaces in $\mathcal{U}_1/D$ and

$$\text{DisCoord}^{\mathcal{U}_1/D}([A_1]_D, [B_1]_D, [C_1]_D) = 0.$$

Consequently,

$$\text{DisCoord}^{\mathcal{U}/D}([A]_D, [B]_D, [C]_D) = \text{DisCoord}(A, B, C).$$

□

3.6 Linear Information Theory Equalities

Most work in Chapter 3 up until this point has been in defining discoordination, showing properties of this quantity, and proving Theorems 3.1.4 and 3.1.8.

In this section we drive equalities involving the discoordination of three subspaces that will be used in our approach to the coded caching problem. We will show Corollary 3.1.7. In all the following equalities let U_1, U_2, U_3 be three subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$. For ease of notation, let A, B, C denote U_1, U_2, U_3 respectively.

$$\begin{aligned} \text{DisCoord}(A, B, C) = &\dim\big(C \cap (A + B)\big) \\ &- \dim(C \cap A) - \dim(C \cap B) \\ &+ \dim(A \cap B \cap C); \end{aligned} \tag{3.6.1}$$

Proof of equality (3.6.1). By Theorem 3.1.4, we can decompose $\mathcal{U}$ into $\mathcal{U}_1, \mathcal{U}_2$ such that A, B, C factor through this decomposition and

1. $A \cap \mathcal{U}_1, B \cap \mathcal{U}_1, C \cap \mathcal{U}_1$ are coordinated in $\mathcal{U}_1$, and
2. there is an isomorphism $\mu : \mathcal{U}_2 \to \mathbb{F}^2 \times \mathbb{F}^m$ which takes $A \cap \mathcal{U}_2, B \cap \mathcal{U}_2, C \cap \mathcal{U}_2$, respectively to

$$\text{Span}(e_1) \otimes \mathbb{F}^m,\ \text{Span}(e_2) \otimes \mathbb{F}^m,\ \text{Span}(e_1 + e_2) \otimes \mathbb{F}^m.$$

It suffices to show this equality holds for the decomposition of A, B, C in $\mathcal{U}_1$ and in $\mathcal{U}_2$. Consider $A_2 = A \cap \mathcal{U}_2, B_2 = B \cap \mathcal{U}_2, C_2 = C \cap \mathcal{U}_2$ under μ, then we have

$$\begin{aligned}
&\text{DisCoord}(A_2, B_2, C_2) = m,\\
&\dim\big(C_2 \cap (A_2 + B_2)\big) = \dim\big(\text{Span}(e_1 + e_2) \otimes \mathbb{F}^m\big) = m,\\
&\dim(C_2 \cap A_2) = \dim(C_2 \cap B_2) = \dim(A_2 \cap B_2 \cap C_2) = 0.
\end{aligned}$$

Thus, (3.6.1) holds for A_2, B_2, C_2.

Consider $A_1 = A \cap \mathcal{U}_1$, $B_1 = B \cap \mathcal{U}_1$, $C_1 = C \cap \mathcal{U}_1$, from Theorem 3.1.9 and since $\text{DisCoord}(A_1, B_1, C_1) = 0$ we have that

$$\begin{aligned}
\dim(A_1) + \dim(B_1) + \dim(C_1) = &\dim(A_1 \cap B_1 \cap C_1)\\
&+ \dim\big((A_1 \cap B_1) + (A_1 \cap C_1) + (B_1 \cap C_1)\big)\\
&+ \dim(A_1 + B_1 + C_1),
\end{aligned} \tag{3.6.2}$$

from the dimension formula we get the following equalities

$$\begin{aligned}
&\dim(A_1 + B_1 + C_1) = \dim(A_1 + B_1) + \dim(C_1) - \dim\big((A_1 + B_1) \cap C_1\big)\\
&= \dim(A_1) + \dim(B_1) + \dim(C_1) - \dim(A_1 \cap B_1) - \dim\big((A_1 + B_1) \cap C_1\big),
\end{aligned}$$

and[3]

$$\begin{aligned}
&\dim\big((A_1 \cap B_1) + (A_1 \cap C_1) + (B_1 \cap C_1)\big) =\\
&\dim(A_1 \cap B_1) + \dim\big((A_1 \cap C_1) + (B_1 \cap C_1)\big) - \dim(A_1 \cap B_1 \cap C_1).
\end{aligned}$$

[3]Since $A_1 \cap B_1 \cap C_1 \subset (A_1 \cap C_1) + (B_1 \cap C_1)$ and as we know $(A_1 \cap C_1) + (B_1 \cap C_1) \subset C_1$, then

$$\big((A_1 \cap C_1) + (B_1 \cap C_1)\big) \cap (A_1 \cap B_1) \subset A_1 \cap B_1 \cap C_1,$$

as a consequence

$$\dim\Big(\big((A_1 \cap C_1) + (B_1 \cap C_1)\big) \cap (A_1 \cap B_1)\Big) = \dim(A_1 \cap B_1 \cap C_1),$$

Substituting these into (3.6.2) we get

$$\dim\big((A_1+B_1)\cap C_1\big)=\dim\big((A_1\cap C_1)+(B_1\cap C_1)\big).$$

Using the above equality the right-hand-side of (3.6.1) can be written as

$$\dim\big((A_1\cap C_1\big)+(B_1\cap C_1))-\dim(A_1\cap C_1)-\dim(B_1\cap C_1)+\dim(A_1\cap B_1\cap C_1).$$

By applying the dimension formula to $A_1\cap C_1$ and $B_1\cap C_1$ we find that the above expression is zero. Then (3.6.1) holds for A_1,B_1,C_1. □

Now we show the following equality from Corollary 3.1.7:

$$\begin{aligned}\mathrm{DisCoord}(A,B,C)=&\dim((A+C)\cap(B+C))-\dim(C)\\&+\dim(A\cap B\cap C)-\dim(A\cap B);\end{aligned}\tag{3.6.3}$$

Proof of equality (3.6.3). Similar to the argument for (3.6.1), it suffices to show this equality holds for $A_1=A\cap\mathcal{U}_1$, $B_1=B\cap\mathcal{U}_1$, $C_1=C\cap\mathcal{U}_1$ and for $A_2=A\cap\mathcal{U}_2$, $B_2=B\cap\mathcal{U}_2$, $C_2=C\cap\mathcal{U}_2$.

Consider A_2,B_2,C_2 under μ (from μ's definition in Theorem 3.1.4) then we have

$$\begin{aligned}&\mathrm{DisCoord}(A_2,B_2,C_2)=m,\\&\dim\big((A_2+C_2)\cap(B_2+C_2)\big)=\dim\big(\mathrm{Span}(e_1,e_2)\otimes\mathbb{F}^m\big)=2m,\\&\dim(C_2)=\dim\big(\mathrm{Span}(e_1+e_2)\otimes\mathbb{F}^m\big)=m,\\&\dim(A_2\cap B_2)=\dim(A_2\cap B_2\cap C_2)=0.\end{aligned}$$

Thus, (3.6.3) holds for A_2,B_2,C_2.

Since $\mathrm{DisCoord}(A_1,B_1,C_1)=0$ we have

$$\dim(A_1\cap B_1\cap C_1)=\dim(C_1\cap A_1)+\dim(C_1\cap B_1)-\dim\big(C_1\cap(A_1+B_1)\big).$$

From the dimension formula we get the following two equalities:

$$\begin{aligned}\dim\big((A_1+C_1)\cap(B_1+C_1)\big) &= \dim(A_1+C_1)+\dim(B_1+C_1)-\dim(A_1+B_1+C_1)\\ &= \dim(A_1)+\dim(B_1)+2\dim(C_1)-\dim(A_1\cap C_1)\\ &\quad -\dim(B_1\cap C_1)-\dim(A_1+B_1+C_1),\end{aligned}$$

and

$$\dim\big(C_1\cap(A_1+B_1)\big) = \dim(C_1)+\dim(A_1+B_1)-\dim(A_1+B_1+C_1).$$

The above qualities show that

$$\begin{aligned}&\dim\big((A_1+C_1)\cap(B_1+C_1)\big)+\dim(A_1\cap B_1\cap C_1)\\ &= \dim(A_1)+\dim(B_1)+\dim(C_1)-\dim(A_1+B_1)\\ &= \dim(C_1)+\dim(A_1\cap B_1)\end{aligned}$$

Which shows that the right-hand-side of (3.6.3) equals zero and as a result (3.6.3) holds for A_1, B_1, C_1. □

Now we show the discoordination and mutual information equality from Corollary 3.1.7.

$$\text{DisCoord}(A,B,C) = \dim(A\cap B\cap C) - I(A;B;C) \tag{3.6.4}$$

Proof of equality (3.6.4). Since

$$I(A;B;C) = I(C;A)+I(C;B)-I(C;B,A),$$

when A, B, C are considered as linear random variables we have that

$$I(A;B;C) = \dim(C\cap A)+\dim(C\cap B)-\dim\big(C\cap(B+A)\big).$$

From (3.6.1) we get that

$$I(A;B;C) = \dim(A\cap B\cap C) - \text{DisCoord}(A,B,C).$$

□

3.6.1 Equalities in Quotient Spaces

We can rewrite (3.6.3) as

$$\text{DisCoord}(A,B,C) = \dim^{\mathcal{U}/C}([A]_C \cap [B]_C) + \dim(A \cap B \cap C) - \dim(A \cap B). \tag{3.6.5}$$

Proof. Since $C \subset (A+C) \cap (B+C)$, we have that

$$\dim^{\mathcal{U}/C}\Big(\big[(A+C)\cap(B+C)\big]_C\Big) = \dim\big((A+C)\cap(B+C)\big) - \dim(C).$$

Given that C is a subspace of both $(A+C)$ and $(B+C)$ we have

$$\dim^{\mathcal{U}/C}\Big(\big[(A+C)\cap(B+C)\big]_C\Big) = \dim^{\mathcal{U}/C}\big([A+C]_C \cap [B+C]_C\big).$$

Since the image of $A+C$ in $\mathcal{U}/C$ equals the image of A in $\mathcal{U}/C$ and the same holds for the images of $B+C$ and B in $\mathcal{U}/C$,

$$\dim^{\mathcal{U}/C}\big([A+C]_C \cap [B+C]_C\big) = \dim^{\mathcal{U}/C}\big([A]_C \cap [B]_C\big),$$

and (3.6.5) follows. □

Considering the notation first introduced in Subsection 3.1.5 to formalize discoordination in quotient spaces, from Corollary 3.1.7, and equality (3.6.5) we have the following useful discoordination equality.

Corollary 3.6.1. *Let U_1, U_2, U_3, W be four subspaces of an arbitrary $\mathbb{F}$-universe, $\mathcal{U}$. Then*

$$\begin{aligned}\text{DisCoord}^{\mathcal{U}/W}\big([U_1]_W, [U_2]_W, [U_3]_W\big) = &\dim^{\mathcal{U}/(U_3+W)}\big([U_1]_{U_3+W} \cap [U_2]_{U_3+W}\big)\\ &+ \dim^{\mathcal{U}/W}\big([U_1]_W \cap [U_2]_W \cap [U_3]_W\big)\\ &- \dim^{\mathcal{U}/W}\big([U_1]_W \cap [U_2]_W\big).\end{aligned}$$

Chapter 4

Coded Caching

So far, we have discussed and formalized linear information theory and alluded to its expressiveness; we aim to further motivate the study of linear information theory by demonstrating its application to the coded caching problem. In this chapter, we formally define the coded caching problem, state what we believe to be the "easiest" open problem in the coded caching context, review relevant literature with a particular focus on the results of Tian [9], and show how our results in linear information theory come at play.

4.1 Problem Statement

There is extensive literature on the many variations of the *coded caching* problem, beginning with the seminal paper [7]; see [8, 10] for a survey of the literature. We start by describing a mild simplification of Maddah-Ali and Niesen's problem description from [7]. We stick to their notation.

Consider a network with a central server that is connected to K users through a shared error-free communication link; the server broadcasts to all users much like a radio broadcast station sends signals to listeners. There is a library of N independent files, denoted $W_1, \ldots, W_N$, each of size F bits. The whole database of files is available to the server. On the contrary, every cache has a limited storage capacity of MF bits, where M is some rational number and we are interested in the case where $0 < M < N$, so the caches can store some information regarding the

files, but not all N files. Hence, each cache can only have the equivalent of M files available in its local storage. The non-locally available content needed by a cache has to be acquired from the server via the shared broadcast communications link.

There are two phases, in the first phase, each user can examine all NF bits of all the files; in this phase, each user knows that in the second phase they need to obtain one of the N files, but do not know which specific file from beforehand. This phase usually takes place during the off-peak hours of the network operation. The first phase is called the *placement phase*, in it every user fills their cache with parts of the content from the library, meaning for $i \in [K]$, user i can store Z_i, a cache of MF bits where Z_i is dependent on $W_1, \ldots, W_N$.

Between the first and the second phase, each user becomes aware of which document they will need; specifically user i specifies a value $d_i \in [N]$; we refer to $\mathbf{d} = (d_1, \ldots, d_K)$ as the *demand vector*. The server is aware of the local content of all the caches.

In the second phase, the *delivery phase*, the Z_i do not change. Delivery takes place during the time when the network is congested. In the delivery phase, the central server is given the demand vector, $\mathbf{d} \in \{1, \ldots, N\}^K$, and the server broadcasts a message $X_{\mathbf{d}}$ conditioned on the cache contents and the demand vector. Every cache needs to reconstruct the file its corresponding user has requested using both the content it has available locally and the message sent by the server.

By a *memory-rate pair* we mean a pair (M, R) of rational numbers; we say such a pair is *achievable* for a given value of (N, K) if for some $F \in \mathbb{N}$, there is a choice of $Z_1, \ldots, Z_K$, such that for all $\mathbf{d} \in \{1, \ldots, N\}^K$ there exists an $X_{\mathbf{d}}$ of at most RF bits such that for all $i \in [K]$, $X_{\mathbf{d}}$ and Z_i determine W_{d_i} (in other words, user i is able to reconstruct file W_{d_i} given Z_i and $X_{\mathbf{d}}$). For a fixed cache size M, the smallest rate R for which (M, R) is achievable characterizes the *memory-rate tradeoff*. Our focus is the problem of completely characterizing the memory-rate tradeoff for a given value of (N, K), or when this is not possible, giving a tight lower bound on the memory-rate tradeoff.

Example 4.1.1. Let $N = K = 2$; this case was solved in [7] and illustrates the main idea. Let $F = 2$, and let $W_1 = (A_1, A_2)$ and $W_2 = (B_1, B_2)$ where $A_1, A_2, B_1, B_2 \in \{0, 1\}$. We claim that the pair $(M, R) = (1/2, 1)$ is achievable: indeed, let $Z_1 =$

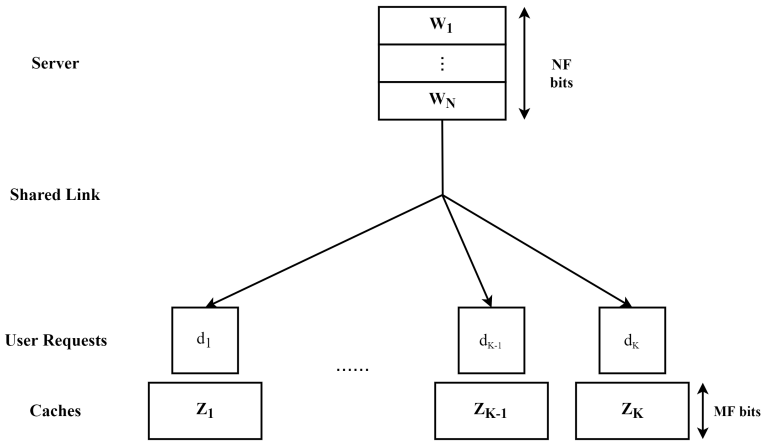

Figure 4.1: Schematic of the coded caching problem.

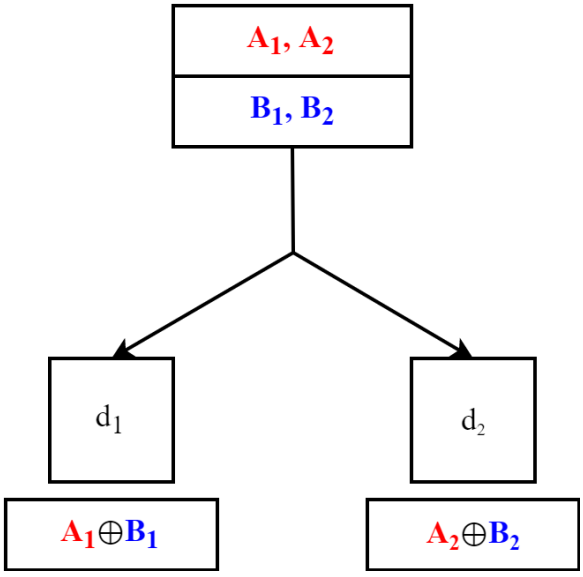

Figure 4.2: Schematic for Example 4.1.1

$A_1 \oplus B_1$ and $Z_2 = A_2 \oplus B_2$. Figure 4.2 demonstrates the caching scheme for this case. For ease of notation, let us denote $X_{(d_1,\ldots,d_K)}$ by $X_{d_1\ldots d_K}$. If $\mathbf{d} = (1,1)$, i.e., both users want W_1, we set $X_{11} = W_1$, i.e., in the delivery phase the server broadcasts $W_1 = (A_1, A_2)$. Similarly we may take $X_{22} = W_2$. If $\mathbf{d} = (1,2)$, i.e., user 1 wants W_1 and user 2 wants W_2, we see that we may take $X_{12} = (A_2, B_1)$, so that X_{12} and Z_1 allow user 1 to determine (A_1, A_2) and user 2 to determine (B_1, B_2). Similarly we can take $X_{21} = (A_1, B_2)$. Hence each cache Z_i stores $MF = 1$ bit, and each $X_{\mathbf{d}}$ can consist of only $RF = 2$ bits.

Let us make a few important remarks about the coded caching problem described. First, if the memory-rate pair (M,R) is achievable for documents $W_1,\ldots,W_N$, of size F, then it is also achievable for documents $W'_1,\ldots,W'_N$ of size tF for any $t \in \mathbb{N}$. Second, if (M_1,R_1) and (M_2,R_2) are achievable for fixed (N,K) then so is any convex combination of these points, i.e., for any rational $0 \leq \alpha \leq 1$,

$$\alpha(M_1,R_1) + (1-\alpha)(M_2,R_2).$$

Third, it is simpler to allow (M,R) to be a pair of real numbers rather than rational numbers; to do this we say that (M,R) is achievable if some limit point of (M,R) is achievable. Fourth, there are some trivial lower bounds on (M,R); for example,

$$R + KM \geq K, \quad NR + M \geq N,$$

which follow from the fact that $X_{(1,2,\ldots,K)}$ and $Z_1, Z_2, \ldots, Z_K$ determines $W_1,\ldots,W_K$,[1] and $X_{(1,\ldots,1)}, X_{(2,\ldots,2)}, \ldots, X_{(N,\ldots,N)}$ and Z_1 determines $W_1,\ldots,W_N$. Fifth, in the original Maddah-Ali and Niesen definition [7], the Z_i are allowed to be random functions of the N documents rather than deterministic functions; the point is that, using Fano's inequality, any lower bound that uses information theory will also produce a lower bound under their assumptions (if the Z_i and $X_{\mathbf{d}}$ are random beyond the values of the documents). Sixth, the coded caching problem can be stated purely information theoretically with the following statements:

- $H(W_i) = F$ for $i \in [N]$, and $H(W_1,\ldots,W_N) = NF$.

[1] Here it is implied that $K \leq N$. If $K > N$ we have the bound $R + NM \geq N$.

- $H(Z_i) \leq MF$ and $H(X_{\mathbf{d}}) \leq RF$ for $i \in [K]$, $\mathbf{d} \in [N]^K$.
- $I(X_{\mathbf{d}}, Z_i; W_{d_i}) = H(W_{d_i}) = F$ for $i \in [K]$ and $\mathbf{d} \in [N]^K$ such that $\mathbf{d} = (d_1, \ldots, d_K)$.

4.2 Relevant Literature

Prior to our work, the strongest memory-rate tradeoff lower bounds were based on a few ideas of [7], and one principle of [10] that can be understood as a "little birdie" principle.

The original [7] paper on coded caching studied and completely characterized the memory-rate tradeoff for the $N = K = 2$ case, by showing the following bounds:

$$R + 2M \geq 2, \quad 2R + M \geq 2, \quad 2R + 2M \geq 3.$$

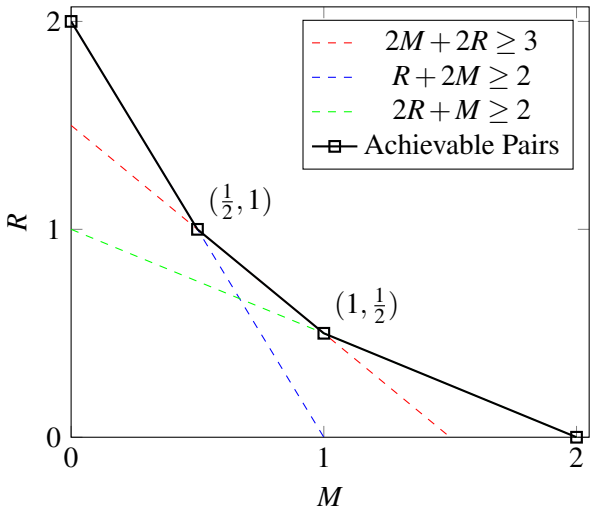

Figure 4.3: Memory-rate tradeoff for the $N = K = 2$ case.

The two trivial bounds, $R + 2M \geq 2$ and $2R + M \geq 2$ are obtained by considering the fact that (Z_1, Z_2, X_{12}) implies (W_1, W_2) and that (Z_1, X_{11}, X_{22}) implies (W_1, W_2). The remarkable bound of [7] is the bound $2R + 2M \geq 3$, obtained by the following

argument: they observe that

$$2MF + 2RF \geq H(X_{12}Z_1) + H(X_{21}Z_2)$$

where H denotes the usual entropy of a random variable (H_2 from Section 2.8), and then use

$$\begin{aligned} H(X_{12}Z_1) + H(X_{21}Z_2) &= H(X_{12}Z_1X_{21}Z_2) + I(X_{12}Z_1; X_{21}Z_2) \\ &\geq H(W_1W_2) + H(W_1) = 2F + F = 3F. \end{aligned}$$

We will derive a bound which is a refinement of this idea, where we use linear information theory, in particular Corollary 3.1.7, to get new memory-rate lower bounds for the $N = K = 3$ case.

There have been many improvements to the original methods of Maddah-Ali and Niesen [3, 9]. One bound of [10] subsumes many of them and can be viewed as a "little birdie" principle or an application of "genie-aided" outer bounds in information theory. The little birdie principle is in effect the following bound (Lemma 2 in [10]):

$$H\big(X_{(d_1,\ldots,d_K)}\big) \geq \sum_{i=1}^{min\{N,K\}} H\big(W_{d_i} \mid Z_1Z_2\ldots Z_i\, W_{d_1}W_{d_2}\ldots W_{d_{i-1}}\big).$$

In the case where $N = K = 3$ the bound becomes

$$H(X_{123}) \geq H(W_1 \mid Z_1) + H(W_2 \mid Z_1Z_2W_1) + H(W_3 \mid Z_1Z_2Z_3W_1W_2).$$

With the assumption that the caches are ordered, suppose there is a "little birdie" (or a "genie") that gives cache i the contents of all the previous caches and the files they reconstructed. Referring to the inequality above, the right-hand-side is what user 1 needs to reconstruct W_1, plus what user 2 needs to reconstruct W_2 given knowledge of the cache and the reconstructed requested file of user 1 plus a similar term for user 3. Nearly all bounds in [10] are dependent on the little birdie bound and

$$H(X_{\mathbf{d}}) + H(Z_i) \geq H(X_{\mathbf{d}}, Z_i) \geq H(W_{d_i}) + H(X_{\mathbf{d}}, Z_i \mid W_{d_i}),$$

where $i \in [K]$ and $\mathbf{d} \in [N]^K$ such that $\mathbf{d} = (d_1, \ldots, d_K)$.

Theorems 1 and 2 of [10] give lower bounds on the optimal rate for any N and K; for instance Theorem 2 gives the bound $2M + 3R \geq 5$ for $N = K = 3$. They also show that for large enough N and $K \leq 5$, theorems 1 and 2 completely characterize the memory-rate tradeoff (see Remark 7 of [10]).

In [9], Tian showed the exact memory-rate tradeoff when there are only two users with an induction proof based on a hypothesis formed by his computer-aided approach. Theorem 4.5 in [9] shows the memory-rate tradeoff of $K = 2$ and $N \geq 3$ is completely characterized with the following two bounds:

$$3M + NR \geq 2N, \quad M + NR \geq N.$$

Here, $M + NR \geq N$ is the trivial bound shown in [7] and $3M + NR \geq 2N$ is a new result from [9].

Remarkably, Tian completely characterized the memory-rate tradeoff for $N = 2$ and $K = 3$ by showing the non-trivial bound $3M + 3R \geq 5$. Meaning, in both cases, $K = 2, N \geq 2$ and $K = 3, N = 2$, the memory-rate tradeoff is completely characterized.

Tian's work also shows that $K = N = 3$ and $N = 2, K = 4$ both require at least 3 non-trivial bounds for their complete memory-rate characterization. Furthermore, the currently known bounds for $K = N = 3$ are tight everywhere except $1/3 \leq M \leq 1$ and the bounds known for $N = 2, K = 4$, are tight everywhere except $1/4 \leq M \leq 2/3$. We suspect the easiest[2] open problem is likely $(N, K) = (3, 3)$.

4.3 The Methods of Tian for $N = K = 3$

Let us refer to [9] by referring to the work's author, Dr. Chao Tian. As stated in Section 4.2, Tian generated many memory-rate lower bounds using a computer aided search through linear programs based on the information of collections of Z_i, $X_{\mathbf{d}}$, and W_i. Let us focus on the case $N = K = 3$, which is—according to the methods of Tian—likely the "easiest" open case of coded caching. Previous to

[2] One can argue that $(N, K) = (2, 4)$ is the "easiest" open problem, however given the work done in [3] for coded caching when N and K are equal, the $N = K = 3$ case is more studied a more intriguing instance of the code caching problem.

Tian's work, the optimal value of R was known for all M except $1/3 \leq M \leq 1$. The fact that $(M,R) = (1,1)$ is attainable was proven by [7], who showed that

$$3R+M \geq 3,\ 3R+2M \geq 5, R+3M \geq 3,$$

and that these lower bounds are tight for $M \geq 1$, since $(M,R) = (1,1),(2,1/3),(3,0)$ are achievable. It was shown that $(M,R) = (1/3,2)$ is achievable by [3], which settles the $M \leq 1/3$ case via $3R+M \geq 3$. Tian shows that for $1/3 \leq M \leq 1$ we have two inequalities

$$M+R \geq 2, \quad 2M+R \geq 8/3;$$

the intersection point of these lines is $(M,R) = (2/3,4/3)$, and Tian shows that such a scheme cannot be given with all the Z_i and $X_{\mathbf{d}}$ being linear functions of the bits of W_1, W_2, W_3.

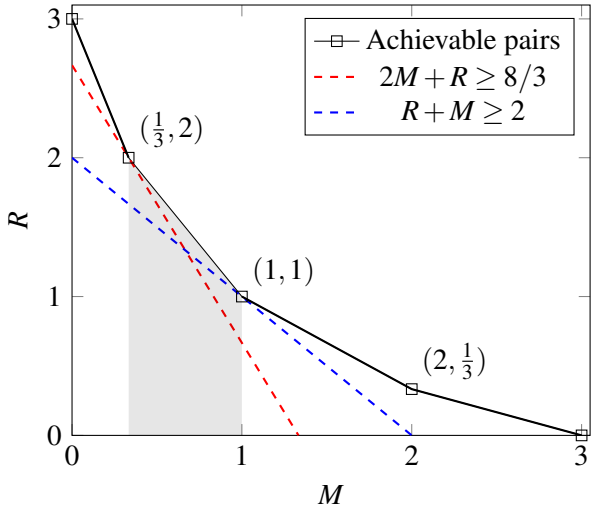

Figure 4.4: Memory-rate tradeoff for the $N = K = 3$ case (without our results). Note that the memory-rate tradeoff is fully characterized for all M except $\frac{1}{3} \leq M \leq 1$ (the shaded region).

Tian's linear program shows that with $m = 1/3 = M/2$ (Table 4 in [9]), we

have

$$H(Z_1 \mid W_1) = 2m, \quad H(Z_1 \mid W_1W_2) = m,$$

$$H(Z_1Z_2 \mid W_1W_2) = 2m, \quad H(Z_1Z_2Z_3 \mid W_1W_2) = 3m.$$

This allows him to conclude (see discussion below Table 4 in [9]), if all random variables are linear functions of the bits of W_1, W_2, W_3, then F must be divisible by 3 (if we are to achieve $(2/3, 4/3)$ with a given value of F), hence we can subdivide the bits of each W_i into three groups,

$$W_1 = A_1A_2A_3, \quad W_2 = B_1B_2B_3, \quad W_3 = C_1C_2C_3, \tag{4.3.1}$$

each group of size $F/3$ (in some cases we might denote the division of W_i into 3 parts by $W_{i1}W_{i2}W_{i3}$ for ease of notation), and we have

$$Z_i = \mathcal{L}_i(A_i, B_i, C_i) \tag{4.3.2}$$

where $\mathcal{L}_i$ is some linear function. Tian gives an argument to show that no such $\mathcal{L}_i$ can achieve $(M,R) = (2/3, 4/3)$. In fact, Tian's argument shows that under a certain subset of the above conditions, we must have $2R + 3M \geq 5$. Let us make this precise.

Definition 4.3.1. Let *Linear Coded Caching* be when the caches and server responses are linear codes of the files; meaning when we constrain the coded caching problem to linear codes. This means all the random variables in the problem (i.e., the Z_i, W_j, and $X_{\mathbf{d}}$) are classical linear random variables (see Definition 2.8.1). Further, let the random variables represent their corresponding linear random variable (see Definition 2.8.4) which can be considered as a linear subspace of some $\mathbb{F}$-universe where $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$.

Definition 4.3.2. Consider the case of coded caching for $N = K = 3$, where F is divisible by three, and each files is subdivided in groups of $F/3$ bits as in (4.3.1). We say that Z_1, Z_2, Z_3 are *separated* (respectively, *separated linear*) if we have (4.3.2) for each $i \in [3]$, for some function (respectively, linear function) $\mathcal{L}_i$.

Tian's argument in Section 5.4 of [9], in effect, proves the following theorem.

Theorem 4.3.3. *For the linear coded caching problem with $N = K = 3$ and F divisible by 3, let the bits of W_1, W_2, W_3 be subdivided into groups of $F/3$ bits as in* (4.3.1)*. If the Z_i are separated linear and*

$$Z_i = (A_i \oplus B_i,\ B_i \oplus C_i)$$

for all $i \in [3]$, then[3]

$$2R + 3M \geq 5.$$

Moreover, if $R'F = \dim(X_{123})$, then $2R' + 3M \geq 5$, and similarly with the indices $1,2,3$ permuted in any way.

Proof. The dimension of X_{123} is $R'F$ for some $R' \leq R$; we will show that

$$2R' + 3M \geq 5.$$

Let us introduce coordinates on $W = W_1 + W_2 + W_3$ so that each element of W, and therefore of X_{123} (and Z_1, Z_2, Z_3), is associated with a vector of $3F$ scalars (i.e. an element of $\mathbb{F}^{3F}$ with $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$).

Choose an arbitrary basis, $\mathcal{A}_1$, of A_1, and similarly bases choose $\mathcal{A}_2, \ldots, \mathcal{C}_3$ of $A_2, \ldots, C_3$. Hence each basis contains $F/3$ elements of some W_i, and we let $\mathcal{W}$ be the union of these bases, meaning $\mathcal{W} = \mathcal{A}_1 \cup \cdots \cup \mathcal{C}_3$.

If $u \in W$, we use $\iota_{\mathcal{W}}(u)$, or simply $\iota(u)$, to denote the element of $\mathbb{F}^{3F}$ associated with u in the coordinates $\mathcal{W}$; then ι can be viewed as an isomorphism $W \to \mathbb{F}^{3F}$. It will be useful to describe vectors in $\mathbb{F}^{3F}$ as blocks of 9 vectors (and similarly for matrices each of whose rows are vectors in $\mathbb{F}^{3F}$); in this case we have ordered $\mathcal{W}$ as

$$\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$$

(the order of the basis elements in each block $\mathcal{A}_1, \ldots, \mathcal{C}_3$ is unimportant).

We have that $\iota(X_{123})$ is a subspace of $\mathbb{F}^{3F}$; choosing an arbitrary basis of X_{123}

[3] This theorem holds for any separated linear caching scheme in which $Z_i, X_{\mathbf{d}}$ implies W_{d_i} and $\{W_{ji} \mid j \in [N] \text{ such that } j \neq d_j\}$. For instance, $Z_1 X_{123}$ implies W_1, B_1, C_1. Another scheme which satisfies these conditions is $Z_i = (A_i, B_i, C_i)$.

and letting G be the matrix whose rows are ι of these basis vectors we get

$$\iota(X_{123}) = \text{RowSpace}(G)$$

(the row space of G) where G is an $R'F \times 3F$ matrix which we view as consisting of 9 blocks

$$G = \begin{bmatrix} G_1 & G_2 & G_3 & \cdots & G_9 \end{bmatrix}.$$

Similarly, by choosing a basis for Z_1, we get

$$\iota(Z_1) = \text{RowSpace}(G')$$

where G' is an $MF \times 3F$ matrix which we view as consisting of 9 blocks

$$G' = \begin{bmatrix} G'_1 & G'_2 & G'_3 & \cdots & G'_9 \end{bmatrix}.$$

It follows that $\iota(X_{123} + Z_1)$ equals the row space of the block matrix

$$\iota(X_{123} + Z_1) = \text{RowSpace}\left(\begin{bmatrix} G' \\ G \end{bmatrix}\right),$$

where

$$\begin{bmatrix} G' \\ G \end{bmatrix} = \begin{bmatrix} G'_1 & G'_2 & G'_3 & G'_4 & G'_5 & G'_6 & G'_7 & G'_8 & G'_9 \\ G_1 & G_2 & G_3 & G_4 & G_5 & G_6 & G_7 & G_8 & G_9 \end{bmatrix}.$$

Consider this matrix with its columns rearranged into two blocks as such

$$\tilde{G} = \left[\begin{array}{ccccc|cccc} G'_1 & G'_2 & G'_3 & G'_4 & G'_7 & G'_5 & G'_6 & G'_8 & G'_9 \\ G_1 & G_2 & G_3 & G_4 & G_7 & G_5 & G_6 & G_8 & G_9 \end{array}\right];$$

we can see that G and $\tilde{G}$ have the same rank. By assumption X_{123} and Z_1 determine A_1, A_2, A_3, B_1, C_1, then $X_{123} + Z_1$ contains each vector of the bases $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{B}_1, \mathcal{C}_1$. Hence, $\iota(X_{123} + Z_1)$ contains each standard basis vector associated the these five bases. By the basis exchange theorem[4], we can apply elementary

[4]We remark that the rows of G' and of G are not necessarily independent, namely if Z_1 and X_{123} have a non-trivial intersection. Still, we can choose a subset of the rows of the matrix formed by the rows of G' and G and apply the basis exchange theorem there. Alternatively, the same argu-

(i.e., invertible) row operations on $\tilde{G}$ to get a matrix

$$\hat{G} = \begin{bmatrix} I & 0 \\ L_1 & L_2 \end{bmatrix},$$

where I is a $5F/3 \times 5F/3$ identity matrix, and 0 is a $5F/3 \times 4F/3$ zero matrix, and L_1, L_2 are some matrices; since the total number of rows of $\tilde{G}$ is at most $MF + R'F$, the number of rows in the L_1, L_2 block matrix is at most $(M + R' - 5/3)F$. Hence the column space of the two rightmost blocks,

$$\text{ColumnSpace}\left(\begin{bmatrix} 0 \\ L_2 \end{bmatrix}\right)$$

is at most $(M + R' - 5/3)F$. The row operations changing $\tilde{G}$ to $\hat{G}$ do not change the column space of any subset of columns of these matrices, then the span of the columns of

$$\begin{bmatrix} G'_5 & G'_6 & G'_8 & G'_9 \\ G_5 & G_6 & G_8 & G_9 \end{bmatrix} \tag{4.3.3}$$

is of dimension at most $(M + R' - 5/3)F$. In particular, the same bound holds for the span of the columns of

$$\begin{bmatrix} G_5 & G_6 & G_8 & G_9 \end{bmatrix}. \tag{4.3.4}$$

By symmetry, using Z_2, the column space of

$$\begin{bmatrix} G_1 & G_3 & G_7 & G_9 \end{bmatrix}$$

has dimension at most $(M + R' - 5/3)F$; using Z_3, the same holds for the column space of

$$\begin{bmatrix} G_1 & G_2 & G_4 & G_5 \end{bmatrix}.$$

Since each column of G appears once or twice in the three block matrices above,

ment works if G' consists only of rows such that the block matrix of rows of G' and G are linearly independent, and the row space of this combined matrix equals $\iota(X_{123} + Z_1)$.

the entire column space of G has dimension at most

$$3(M+R'-5/3)F.$$

But the dimension of the column space of G is the rank of G, which by assumption equals $R'F$, hence

$$R' = \text{Rank}(G) \leq 3(M+R'-5/3)F$$

It follows that $3M+2R' \geq 5$. □

Table 4 in [9] implies that if $(M,R) = (2/3,4/3)$ is achievable by a linear caching scheme, then the Z_i must be as in Theorem 4.3.3. This contradicts the hypothesis of Theorem 4.3.3 and leads Tian to conclude that $(M,R) = (2/3,4/3)$ cannot be achieved by any separated linear coding scheme given that the values of Table 4 in [9] hold exactly.

We remark that Theorem 4.3.3 does not really analyze X_{123} directly, rather it draws conclusions based on the particular nature of the Z_i and the fact that X_{123} and Z_i determine W_i.

4.4 Symmetrization and Averaging

Let us review useful forms of averaging to simplify certain expressions in coded caching.

4.4.1 Symmetrization as Averaging

The symmetric group S_K of permutations on $\{1,\ldots,K\}$ acts on the K users of a coded caching problem, and similarly S_N acts on the N documents. Since these two actions are independent of each other (i.e., can be performed in either order), this gives us an action of $S_K \times S_N$ on all the expressions involving the indices of $W_i, Z_i, X_{\mathbf{d}}$: namely for $\kappa \in S_K$ and $\nu \in S_N$, we set

$$(\kappa,\nu)W_i = W_{\nu(i)}, \quad (\kappa,\nu)Z_i = Z_{\kappa(i)},$$

and

$$(\kappa,\nu)X_{\mathbf{d}} = X_{(\kappa,\nu)\mathbf{d}}, \text{ where } (\kappa,\nu)\mathbf{d} = (\kappa,\nu)(d_1,\ldots,d_K) = \big(\nu(d_{\kappa(1)}),\ldots,\nu(d_{\kappa(K)})\big)$$

(since each d_i represents a value in $[N]$ of a document requested by a user $i \in [K]$).

The action of $S_K \times S_N$ similarly extends to any combination of W's, Z's, and X's. We use $\dim^{\text{avg}}$ to denote the average dimension of any expression under this group action. For example, for $K = N = 3$,

$$\dim^{\text{avg}}(Z_1 + W_1 + W_3 + X_{122}) = \frac{1}{K!\,N!} \sum_{(\kappa,\nu)\in S_K\times S_N} \dim\big(Z_{\kappa(1)} + W_{\nu(1)} + W_{\nu(3)} + X_{(\kappa,\nu)(1,2,2)}\big).$$

Hence we have

$$\dim^{\text{avg}}(Z_1 + W_1 + W_3 + X_{122}) = \dim^{\text{avg}}(\kappa,\nu)(Z_1 + W_1 + W_3 + X_{122})$$

for any $\kappa \in S_K$ and $\nu \in S_N$, for instance

$$\dim^{\text{avg}}(Z_1 + W_1 + W_3 + X_{122}) = \dim^{\text{avg}}(Z_3 + W_1 + W_2 + X_{331}).$$

We can use he same idea to define $\text{DisCoord}^{\text{avg}}(W_i, W_j, Z_k)$ for any $i, j \in [N]$ and any $k \in [K]$ in the context of the linear coded caching problem. More precisely, we have

$$\text{DisCoord}^{\text{avg}}(W_i, W_j, Z_k) = \frac{1}{K!\,N!} \sum_{(\kappa,\nu)\in S_K\times S_N} \text{DisCoord}\big(W_{\nu(i)}, W_{\nu(j)}, Z_{\kappa(k)}\big).$$

This averaging technique is convenient in proving lower bounds on achievable memory-rate pairs. See [10], for particular uses of this technique; we will use it in our bounds as well.

4.4.2 Symmetric Coded Caching Schemes

In [9], Tian prefers to symmetrize the coded caching schemes beforehand. In other words, for a given coded caching scheme on documents of size F, each element of

$S_K \times S_N$ yields a new scheme of size F, and by concatenating the schemes he gets a new scheme on documents of size $K!N!F$ bits such that the dimension of any expression in the W's, Z's, and X's is invariant under the action of $S_K \times S_N$. See Proposition 3 of Section 3.3 of [9].

The same follows for proving lower bounds, it suffices to consider the case where the coded caching scheme has dimensions of all expressions in W, Z, X that are invariant under this $S_K \times S_N$ action (i.e., the case where dim and $\dim^{\text{avg}}$ are equal).

4.4.3 A Lopsided Example: Average and Worst Case

To motivate symmetrization, we give the following example of a "highly non-symmetric" (or "lopsided") scheme with $N = K = 3$ where X_{123} can be taken to be 0. Here we are assuming $X_{\mathbf{d}}$ and Z_i are linear functions of W_1, W_2, W_3.

Consider the case $M = 1$ where we set $Z_i = W_i$ for all $i \in [3]$. In this case we can take $X_{123} = 0$. While Theorem 4.3.3 shows that X_{123} has dimension $R'F$ with $R' \geq (5-3M)/2$, the same cannot be said of this particular scheme as it does not follow the conditions of theorem. To show $2R + 3M \geq 5$ holds we need to prove that some X_{ijk} must have dimension at least $(5-3M)/2$. We remark that if we use symmetrization and somehow prove that some X_{ijk} has dimension at least $(5-3M)/2$ we also prove the stronger fact that the average dimension of X_{ijk} with i, j, k distinct is at least $(5-3M)/2$.

It is instructive to compare the average versus worst case here. We may take $X_{213} = W_1 \oplus W_2$, so that X_{213} can be of dimension F, and similarly (the two other single transpositions) X_{321} and X_{132} can be taken to have dimension F. However, we claim that that X_{312} must be of dimension at least $2F$ under this scheme: indeed, for X_{312} and Z_1 to determine W_3, X_{312} must contain $W_3 + \mathcal{L}_1(Z_1)$ for some linear map $\mathcal{L}_1$, and similarly X_{312} must contain $W_1 + \mathcal{L}_2(Z_2)$ and $W_2 + \mathcal{L}_3(Z_3)$ for linear maps $\mathcal{L}_2, \mathcal{L}_3$ to determine W_1 and W_2 given Z_2 and Z_3 respectively. Therefore, X_{123} has the same row space as a matrix of the form

$$\begin{bmatrix} L_1 & 0 & I \\ I & L_2 & 0 \\ 0 & I & L_3 \end{bmatrix},$$

where 0 and I are respectively zero and identity matrices of size $F \times F$ and the L_i are $F \times F$ matrices corresponding to the $\mathcal{L}_i$. Then, X_{312} has the same row space as the above $3F$ by $3F$ matrix. By dropping the first row and last column, we see that X_{123} has rank at least that of

$$\begin{bmatrix} I & L_2 \\ 0 & I \end{bmatrix}.$$

We can eliminate the L_2 from the above matrix with row operations, leaving an identity matrix of size $2F \times 2F$. Consequently, the dimension of X_{312} must be at least $2F$ (and this suffices, since we can verify that setting X_{312} to be $W_1 \oplus W_2, W_1 \oplus W_3$ satisfies the conditions of each user). A similar calculation holds for X_{231} which is the other full cycle permutation.

Hence, under the lopsided scheme $Z_i = W_i$, the maximum dimension of an X_{ijk} is $2F$, and the average over all distinct i, j, k is at least $(3 \cdot F + 2 \cdot 2F)/6 = 7F/6$.

4.5 The Z-Decomposition Lemma

In this section we describe what an individual Z_i must look like in terms of its linear algebra in the linear coded caching problem for $N = K = 3$.

Definition 4.5.1. Let Z be linear subspace of an $\mathbb{F}$-universe, $\mathcal{U}$. Suppose Z has dimension n and $\mathcal{U}$ has a decomposition A, B, C. We say that

1. Z is a *pure individual scheme* if Z is spanned by $A' = Z \cap A$, $B' = Z \cap B$, and $C' = Z \cap C$, in which case $Z = A' + B' + C'$;

2. Z is a *pure Tian scheme*[5] if there exist $A' \subset A$, $B' \subset B$, and $C' \subset C$ such that A', B', C' are of the same dimension, $m = n/2$, and there are bases $a'_1, \ldots, a'_m$ of A', $b'_1, \ldots, b'_m$ of B', and $c'_1, \ldots, c'_m$ of C' such that Z is spanned by $a'_i + b'_i, b'_i + c'_i$ for $i \in [m]$; hence, in the notation of Subsection 2.3.1, we have Z is the span of $A' \oplus_{\mu_1} B'$ and $B' \oplus_{\mu_2} C'$ where μ_1 is the isomorphism $A' \to B'$ taking a'_i to b'_i for all $i \in [m]$, and similarly for $\mu_2 \colon B' \to C'$ taking b'_i to c'_i;

3. Z is a *pure AB-scheme* if there exist $A' \subset A$, $B' \subset B$ such that A', B' are of the same dimension, n, and there are bases $a'_1, \ldots, a'_n$ of A', $b'_1, \ldots, b'_n$ of B' such

[5]The name "Tian scheme" is used here since in [9], Tian stated Theorem 4.3.3 by assuming the caches are of this form (i.e., $Z_i = (A_i \oplus B_i), (B_i \oplus C_i)$ for $i \in [3]$).

that Z is spanned by $a'_i + b'_i$ for $i \in [n]$; hence $Z = A' \oplus_\mu B'$ where μ is the isomorphism $A' \to B'$ taking a'_i to b'_i for all i;

4. we similarly define when Z is a *pure AC-scheme* or a *pure BC-scheme*; and

5. Z is a *pure triple sum scheme* if there exist $A' \subset A$, $B' \subset B$, and $C' \subset C$ such that A', B', C' are of the same dimension, n, and there are bases $a'_1, \ldots, a'_n$ of A', $b'_1, \ldots, b'_n$ of B', and $c'_1, \ldots, c'_n$ of C' such that Z is spanned by $a'_i + b'_i + c'_i$ for $i \in [n]$; in this case $Z = A' \oplus_{\mu_1} B' \oplus_{\mu_2} C'$, where μ_1, μ_2 are, respectively, the isomorphisms $A' \to B'$ and $A' \to C'$ taking a'_i to, respectively, b'_i and c'_i for all i.

Lemma 4.5.2. *Let Z be linear subspace of an $\mathbb{F}$-universe, $\mathcal{U}$, where $\mathcal{U}$ has a decomposition A, B, C. Then there exist subspaces A^j, B^j, C^j indexed on integers $j \in [5]$ such that*

1. *$A^1, \ldots, A^5$ are linearly independent subspaces of A, as are $B^1, \ldots, B^5 \subset B$ and $C^1, \ldots, C^5 \subset C$;*

2. *Z is spanned by:*

   - *$A^1 + B^1 + C^1$ (i.e., an individual scheme);*
   - *$A^2 \oplus B^2$, $B^2 \oplus C^2$ (i.e., a Tian scheme);*
   - *$A^3 \oplus B^3$, $A^4 \oplus C^3$, $B^4 \oplus C^4$ (i.e., an AB-, AC-, and a BC-scheme[6]); and*
   - *$A^5 \oplus B^5 \oplus C^5$ (i.e., a triple scheme).*

Our intended application is to caches, Z_i, in the linear coded caching problem for the $N = K = 3$ case. This lemma shows that any cache is really some combination of the caching schemes in Definition 4.5.1. In a symmetrized scheme, the dimensions of all subspaces with superscripts 3 and 4 are equal.

Our proof is quite straightforward, although somewhat tedious. The strategy is to define the spaces in the following order: A^1, B^1, C^1, then A^2, B^2, C^2, then $A^3, B^3, A^4, C^3, B^4, C^4$, and then A^5, B^5, C^5. In each stage we make the necessary definitions and then show a number of properties of these spaces. Ultimately we

[6]When grouping the *AB*-, *AC*-, and a *BC*-schemes together with the condition that they are of the same size, i.e., they are symmetrized, we might refer to them as a *symmetric two-way scheme*.

need to show that the $A^1, \ldots, A^5$ are linearly independent, similarly show the linear independence of the B^i's and C^i's. Then we need to decompose any $a+b+c \in Z$ with $a \in A$, $b \in B$, and $c \in C$ as a unique sum of the above schemes; the uniqueness is immediate from the linear independence of these subspaces.

Proof. Set $A^1 = Z \cap A$, $B^1 = Z \cap B$, and $C^1 = Z \cap C$.

Say that an $u \in \mathcal{U}$ is *B-pairable* if for some $b \in B$ we have $u+b \in Z$. Let us show that if u is B-pairable with b, then

$$u \text{ is } B\text{-pairable with } b' \in B \iff b = b' + b^1 \text{ for some } b^1 \in B^1 = Z \cap B; \quad (4.5.1)$$

"$\Rightarrow$" follows from the fact that if $u+b$ and $u+b'$ both lie in Z, and then so does $(u+b)-(u+b')$; hence $b+b' \in Z$; since $b, b' \in B$, we have $b+b' \in Z \cap B$.
"$\Leftarrow$" follows from the fact that if $b = b' + b^1$ with $b^1 \in B^1$, then b' lies in B. Since $u+b = (u+b') + (b^1) \in Z$, for $b^1 \in Z$, then $u+b' \in Z$.

We similarly define *A-pairable* and *C-pairable*. A similar remark holds for an element $a \in A$ that pairs with some $u \in \mathcal{U}$, and for a $c \in C$ that pairs with some $u \in \mathcal{U}$.

The set of elements of A that are B-pairable is a subspace $A' \subset A$ which contains all $a \in A^1$ (taking $b = 0$). Similarly, let the subspace $A'' \subset A$ be the set of elements of A that are C-pairable. Then $A' \cap A''$ contains A^1; let $a_1^2, \ldots, a_d^2$ be a basis of $A' \cap A''$ relative to A^1. For any $i \in [d]$, choose a $b_i^2 \in B$ and a $c_i^2 \in C$ such that $a_i^2 + b_i^2$ and $a_i^2 + c_i^2$ lie in Z.

We claim that the $b_1^2, \ldots, b_d^2$ are linearly independent in B/B^1: if not then some linear combination of the b_i^2 lies in B^1, and hence the corresponding linear combination of the a_i^2, say a, has $a+0 \in Z$; but then $a \in Z$, and so $a \in A^1$, contradicting the fact that $a_1^2, \ldots, a_d^2$ is a basis relative to A^1. Similarly the $c_1^2, \ldots, c_d^2$ are linearly independent in C/C^1.

Let A^2 be the span of the a_i^2. Similarly let B^2 and C^2 be the span of the b_i^2 and c_i^2 respectively. Then B^2 is linearly independent from B^1, by the above argument, and similarly C^2 is linearly independent from C^1; by definition A^2 is linearly independent from A^1.

By the definition of A', A'', and A^1, an element $a \in A$ is both B- and C-pairable iff $a \in A' \cap A'' = A^1 + A^2$. Let us prove the analogous statement holds with A, B, C

exchanged: e.g., let us prove that if $b \in B$ is both A- and C-pairable, then $b \in B^1 + B^2$. For any such b there are $a \in A$ and $c \in C$ such that $b+a, b+c, a+c \in Z$; hence $a = a^1 + a^2$ with $a^i \in A^i$ for $i = 1,2$. Then there exists $b^2 \in B^2$ paired with a^2, and $a + b^2 = (a^2 + b^2) + a^1 \in Z$. Thus a is B-paired with $b^2 \in B$ and with $b \in B$, by (4.5.1), $b = b^2 + b^1$ for some $b^1 \in B^1$, meaning $b \in B^1 + B^2$. Similarly, some $c \in C$ is both A- and B-pairable iff $c \in C^1 + C^2$.

Let us now construct $A^3, A^4, B^3, B^4, C^3, C^4$ with the desired properties. Pick a basis, $a_1^3, \ldots, a_s^3$ of A' relative to $A' \cap A''$, and let A^3 be the span of this relative basis; similarly pick a basis $a_1^4, \ldots, a_t^4$, of A'' relative to $A' \cap A''$ and let A^4 denote the span of this relative basis. By the dimension formula $A^1 + A^2$, A^3, A^4 are linearly independent. For each a_i^3, choose a b_i^3 such that $a_i^3 + b_i^3 \in Z$, and similarly pick a b_i^3 such that $a_i^4 + c_i^3 \in Z$. We claim that the b_i^3 are linearly independent in $B/(B^1 + B^2)$; otherwise we would get a corresponding linear combination of the a_i^3, say a, with $a + 0 \in Z$, contradicting the the fact that $a_1^3, \ldots, a_s^3$ form a relative basis. The same argument with the B's and b's replaced by C's and c's show that the c_i^3 are linearly independent in $C/(C^1 + C^2)$. Let B^3 be the span of the b_i^3, and C^3 the span of the c_i^3.

Let $\tilde{B} \subset B$ denote the subspace of *C-paired* elements of B; clearly, $\tilde{B}$ contains B^1, B^2; let $b_1^4, \ldots, b_p^4$ be a basis of $\tilde{B}$ relative to $B^1 + B^2$. For each $i \in [p]$, choose an element c_i^4 such that $b_i^4 + c_i^4 \in Z$; let B^4 be the span of all b_i^4, and C^4 the span of all c_i^4.

Due to the asymmetry in our definition the following is known about $A^1, \ldots, A^4$:

1. $A^1, \ldots, A^4 \subset A$ are linearly independent;
2. the subspace of A of elements that are B-pairable equals $A^1 + A^2 + A^3$;
3. the subspace of A of elements that are C-pairable equals $A^1 + A^2 + A^4$.

We now wish to prove the analogous claims about the B^j and B, and the C^j and C. Let us start with the B^j and B.

1. B^1, B^2 are linearly independent: shown above.
2. B^1, B^2, B^3 are linearly independent: if not, then we have $b^3 = b^2 + b^1$ for some $b^i \in B^i$ and $i \in [3]$ with b_3 nonzero (since B^1, B^2 are linearly independent). Then there exist $a^2 \in A^2$ and $c^2 \in C^2$ such that $a^2 + b^2, a^2 + c^2, b^2 + c^2$

all lie in Z. Further, there is a nonzero $a^3 \in A^3$ such that $b^3 + a^3 \in Z$. But then $a = a^3$ is both B-pairable (with b^3) and C-pairable with $-c^2$, since

$$a - c^2 = a - c^2 + (b^3 - b^2 - b^1) = (a + b^3) - (c^2 + b^2) + (b^1)$$

and $a + b^3$, $c^2 + b^2$, and b^1 all lie in Z. Given that $a = a^3$ is nonzero B- and C-pairable, but not in $A^1 + A^2 = A' \cap A''$, we have a contradiction since by construction $A' \cap A''$ contains all B- and C-pairable vectors in A.

3. B^1, B^2, B^3, B^4 are linearly independent: if not, then we have $b^4 = b^3 + b^2 + b^1$ for some $b^i \in B^i$ with b^4 nonzero (since B^1, B^2, B^3 are linearly independent). Given that the b_i^4 form a basis of B' relative to B^1, B^2, we have $b^3 \neq 0$. There there exists a linear combination, c^4, of the c_i^4, such that $b^4 + c^4 \in Z$, and similarly exists a $a^3 \in \text{Span}(a_1^3, \ldots, a_s^3)$ such that $b^3 + a^3 \in Z$. Similarly, b^2 has a corresponding a^2 and c^2 such that both $b^2 + a^2$ and $b^2 + c^2$ lie in Z. Since

$$b^4 + c^4,\ b^3 + a^3,\ b^2 + a^2,\ b^2 + c^2,\ b^1$$

all lie in Z, we have that b^3 is A-pairable and

$$b^3 + (c^4 - c^2) = b^4 - b^2 - b^1 + c^4 - c^2 = (b^4 + c^4) - (b^2 + c^2) - b^1 \in Z.$$

Hence if $a = a^3$, $b = b^3$, and $c = c^4 - c^2$, then $a + b \in Z$ and $b + c \in Z$ and hence $a + c \in Z$. Meaning a is both B- and C-pairable, but $a = a^3 \notin A^1 + A^2$, as a^3 is a linear combination of vectors in A^3 alone, which is a contradiction.

4. $b \in B$ is A-pairable iff $b \in B^1 + B^2 + B^3$ (the converse holds by definition of B^1, B^2, B^3): if b is A-pairable with a, then a is B-pairable and hence $a = a^3 + a^2 + a^1$. There exist b^3, b^2 in B^3, B^2 respectively such that $a^i + b^i \in Z$ for $i = 2, 3$. But then a is B-pairable by $b^3 + b^2$, and B-pairable by b, (4.5.1) implies that b equals some element of B^1 plus $b^3 + b^2$, and hence $b \in B^1 + B^2 + B^3$.

5. b is C-pairable iff $b \in B^1 + B^2 + B^4$: follows from the definition of B^1, B^2, and B^4.

Now we show similar claims for the C^j and C.

1. C^1, C^2 are linearly independent: proven above.

2. C^1, C^2, C^3 are linearly independent: one argues just as for B^1, B^2, B^3.

3. $c_1^4, \ldots, c_p^4$ are linearly independent: if not, then some nontrivial linear combination of them is zero, and, b, the corresponding linear combination of b_i^4, has $b + 0 \in Z$. This implies that $B^4 \cap B^1$ is nonzero, which contradicts the independence of B_1 and B_4.

4. C^1, C^2, C^3, C^4 are linearly independent: any nonzero element $c^4 \in C^4$ has a corresponding nonzero $b^4 \in B^4$ such that $c^4 + b^4 \in Z$. Since c^4, c^2, c^1 are B-paired, then $c^3 = c^4 - c^2 - c^1$ is also B-paired, say with b'. By definition any $c^3 \in C^3$ can be A-paired, say with a'; hence

$$c^3 + b',\ c^3 + a',\ a' + b'$$

   all lie in Z, and hence a' can be both B- and C-paired, implying that $a' = a^2 + a^1$ for $a^1 \in A^1$ and $a^2 \in A^2$. Given that a^2 can be C-paired with some $\tilde{c}^2 \in C^2$, we have $(a^2 + \tilde{c}^2) + a_1 \in Z$. Since $c^3 + a'$ and $\tilde{c}^2 + a'$ are both in Z, we have $c^3 + \tilde{c}^2 \in Z$, and $c^3 + \tilde{c}^2 \in Z \cap C = C^1$. As C^1, C^2, C^3 are linearly independent, it follows that $c^3 = 0$ (and $\tilde{c}^2 = 0$), meaning $c^4 = c^2 + c^1$. For the nonzero linear combination of the c_i^4 that give c^4, there is a corresponding linear combination of the b_i^4, $b^4 \in B^4$, such that $b^4 \neq 0$ and $c^4 + b^4 \in Z$. However, since c^2, c^1 are both A- and B-pairable, so is c^4, and for some $b'' \in B$, $a'' \in A$ we have

$$c^4 + b'',\ c^4 + a'',\ b'' + a''$$

   are vectors in Z. Since $b^4 + c^4 \in Z$, then $b^4 + a'' \in Z$. Thus b^4 is A- and C-pairable, and $b^4 \in B^1 + B^2$. But this contradicts the linear independence of the B^i and the fact that $b^4 \neq 0$ for $b^4 \in B^4$.

5. $c \in C$ is A-pairable iff $c \in C^1 + C^2 + C^3$: same proof as for $b \in B^1 + B^2 + B^3$.

6. $c \in C$ is B-pairable iff $c \in C^1 + C^2 + C^4$ (the converse holds by the definition of C^1, C^2, C^4): Suppose $c \in C$ is B-pairable with b, then b is C-pairable and from the claims for the B^i we have that $b = b^4 + b^2 + b^1$ with $b^i \in B^i$ for $i = 1,2,4$ such that there exist $c^4 \in C^4$ and $c^2 \in C^2$ with $b^i + c^i \in Z$ for $i = 2,4$. Therefore, b is pairable with $c^4 + c^2$, and by (4.5.1), we see that c equals $c^4 + c^2$ plus some element of C^1; meaning $c \in C^1 + C^2 + C^4$.

Finally we construct A^5, B^5, C^5: to do so, consider the subset $\tilde{A}$ of $a \in A$ such that $a+b+c \in Z$ for some $b \in B$ and $c \in C$. $\tilde{A}$ is a subspace, and contains $A^1, \ldots, A^4$; let $a_1^5, \ldots, a_q^5$ be a basis of $\tilde{A}$ relative to $A^1 + A^2 + A^3 + A^4$, and for each $i \in [q]$ choose b_i^5 and c_i^5 such that $a_i^5 + b_i^5 + c_i^5 \in Z$. Let A^5, B^5, C^5, respectively, be the spans of the a_i^5, the b_i^5, and the c_i^5. We prove the following claims, all with ideas similar to the ideas above.

1. $A^1, \ldots, A^5$ are linearly independent: immediate from the definition of a_i^5.

2. If $a+b+c \in Z$ for some $a \in A$, $b \in B$, $c \in C$, then $a \in A^1 + \cdots + A^5$: follows from the definition of A^i.

3. If $a+b+c \in Z$ for some $a \in A$, $b \in B$, $c \in C$, and $a = a^1 + \cdots + a^5$ with $a^i \in A^i$ such that $a^5 \neq 0$, then $b \notin B^1 + B^2 + B^3 + B^4$: otherwise $b = b^1 + \cdots + b^4$ with $b^i \in B^i$; in this case we have $c^4 \in C^4$ such that $b^4 + c^4 \in Z$ and $\tilde{a}^i \in A^i$ for $i = 2,3$ with $\tilde{a}^i + b^i \in Z$. Since $a+b+c$ is in Z, then

$$(a^1 + \cdots + a^5) + (b^1 + \cdots + b^4) + c \in Z$$

it follows that

$$(a^1 + \cdots + a^5 - \tilde{a}^2 - \tilde{a}^3) + b^1 + c - c^4 \in Z,$$

and hence

$$(a^1 + \cdots + a^5 - \tilde{a}^2 - \tilde{a}^3) + (c - c^4) \in Z,$$

meaning $a^1 + \cdots + a^5 - \tilde{a}^2 - \tilde{a}^3$ is C-pairable and therefore lies in $A^1 + A^2 + A^4$. But this is impossible since $a^5 \neq 0$ and A^5 is linearly independent from $A^1, \ldots, A^4$.

4. If $a+b+c \in Z$ for some $a \in A$, $b \in B$, $c \in C$, and $a = a^1 + \cdots + a^5$ with $a^i \in A^i$ such that $a^5 \neq 0$, then $c \notin C^1 + C^2 + C^3 + C^4$: can be shown with an argument similar to the one above.

5. The b_i^5 are linearly independent: if not, then for the corresponding linear combination of a_i^5, say a^5, and for the corresponding linear combination of c_i^5, say c^5 with $a^5, c^5 \neq 0$ we have that $a^5 + c^5 \in Z$. But then a^5 is C-pairable and must lie in $A^1 + A^2 + A^4$, which contradicts the independence of $A^1, \ldots, A^5$ shown above.

6. The c_i^5 are linearly independent: can be shown with a similar argument to the one above.

7. $B^1, \ldots, B^5$ are linearly independent: if they are dependent, then we have $b^5 = b^1 + \cdots + b^4$ with $b^i \in B^i$ where at least one of the b^i is nonzero; since $B^1, \ldots, B^4$ are linearly independent, we know that $b^5 \neq 0$. By definition there exist nonzero $a^5 \in A^5$ and nonzero $c^5 \in c^5$ such that $b^5 + a^5 + c^5 \in Z$. Additionally for $i \in [3]$, there exist $a^i \in A^i$ such that $a^i + b^i \in Z$ and there exists $c^4 \in C^4$ such that $c^4 + b^4 \in Z$. Let $a = a^5 - a^3 - a^2$; since $b^5 + a^5 + c^5 \in Z$ we have

$$(b^1 + b^2 + b^3 + b^4) + a^5 + c^5 \in Z.$$

which implies

$$(a^5 - a^2 - a^3) + (c^5 - c^4) = a + (c^5 - c^4) \in Z.$$

This means a is C-pairable and $a \in A^1 + A^2 + A^4$. Since $a^5 \neq 0$, this contradicts the linear independence of $A^1, \ldots, A^5$.

8. The $C^1, \ldots, C^5$ are linearly independent: can be shown with similar argument to the one above.

9. If $a+b+c \in Z$ for some $a \in A$, $b \in B$, $c \in C$, then $b \in B^1 + \cdots + B^5$: we have $a \in A^1 + \cdots + A^5$, by definition of the A^i, thus there exist $a^i \in A^i$ such that $a = a^1 + \cdots + a^5$. It follows that $a^5 + b^5 + c^5 \in Z$ for some nonzero b^5, c^5 in B^5 and C^5, respectively. Further, $a^2 + b^2 \in Z$ for some $b^2 \in B^2$, and

similarly a^3+b^3 and a^4+c^3 are in Z for some $b^3 \in B^3$ and $c^3 \in C^3$. Given that $(a^1+\ldots+a^5)+b+c \in Z$, it follows that

$$(b-b^2-b^3-b^5)+(c-c^3-c^5) \in Z,$$

and hence $b-b^2-b^3-b^5$ is C-pairable, consequently, it lies in $B^1+B^2+B^4$. Given that $B^1,\ldots,B^5$ are linearly independent and $b^5 \neq 0$, we conclude that $b \in B^1+\cdots+B^5$.

10. If $a+b+c \in Z$ for some $a \in A$, $b \in B$, $c \in C$, then $c \in C^1+\cdots+C^5$: can be shown with a similar argument to the one above.

At this point we claim that Z consists precisely of the sums given in the lemma. Namely, say that $a+b+c \in Z$ with $a \in A$, $b \in B$, $c \in C$. Then write a as $a^1+\cdots+a^5$ for some $a^i \in A^i$. Corresponding to $a^5 \in A^5$ there are $b^5 \in B^5$ and $c^5 \in C^5$ such that $a^5+b^5+c^5 \in Z$, there exists $b^2 \in B^2$, $b^3 \in B^3$, and $c^3 \in C^3$ such that

$$a^2+b^2,\ a^3+b^3,\ a^4+c^3$$

all lie in Z. Let $\tilde{b}=b+b^5+b^3$ and $\tilde{c}=c+c^5+c^4$ then we have

$$a^1+\tilde{b}+\tilde{c} \in Z,$$

and hence $\tilde{b}+\tilde{c} \in Z$. It follows that $\tilde{b}$ is C-pairable, which implies that $\tilde{b}=\tilde{b}^1+\tilde{b}^2+\tilde{b}^4$ with $\tilde{b}^i \in B^i$. Then corresponding to $\tilde{b}^2,\tilde{b}^4$ there are $\tilde{c}^2 \in C^2$ and $\tilde{c}^4 \in C^4$ such that $\tilde{b}^i+\tilde{c}^i \in Z$ for $i=2,4$. Let $\hat{b}=\tilde{b}^2+\tilde{b}^4$ and $\hat{c}=\tilde{c}^2+\tilde{c}^4$, now we have $b'=\tilde{b}+\hat{c}=\tilde{b}^1 \in Z \cap B$ and $c'=\tilde{c}+\hat{b}=\tilde{c}^1 \in Z \cap C$. So far we have written

$$a=a^1+\cdots+a^5, \quad b=b'+\tilde{b}^2+b^3+\tilde{b}^4+b^5, \quad c=c'+\tilde{c}^2+c^3+\tilde{c}^4+c^5,$$

(with $b' \in B^1$, $c' \in C^1$, $\tilde{b}^i$ or b^i in B^i, $\tilde{c}^i$ or c^i in C^i, and $a^i \in A^i$). Consequently, we have written any triple $(a,b,c) \in \mathcal{U}$ with $a+b+c \in Z$ as a sum of elements of the schemes in the lemma. Meaning that Z is the sum of the schemes in the lemma. Conversely, any triple (a,b,c) that is the sum of the elements (i.e. $a \in A^1+\ldots+A^5$, etc.) of these schemes must have $a+b+c \in Z$. Since the $A^1,\ldots,A^5$ are linearly

independent, as are the B^i's and C^i's, the decomposition of any such triple (a,b,c) is unique. □

4.6 A New Caching Scheme for $N = K = 3$ and $M = 1/2$

In this section we give a caching scheme for the coded caching problem in the $N = K = 3$ case which shows the memory-rate pair $(M,R) = (1/2,\ 5/3)$ is achievable.

Theorem 4.6.1. *In the coded caching problem with $N = K = 3$ and F divisible by 6, the memory-rate pair $(M,R) = (1/2,\ 5/3)$ is achievable.*

Proof. It suffices to give a caching scheme for $N = K = 3$ and $M = 1/2$ such that the worst case server broadcast size over all possible user demands is $\frac{10}{6}F$.

Similar to (4.3.1), let us partition the W_i into six disjoint parts where

$$W_i = W_{i1}, W_{i2}, W_{i3}\ , \quad W_{i1} = W'_{i1}, W''_{i1} \quad \text{for every } i \in [3].$$

In our scheme the caches will be of the following form:

$$Z_j = W'_{1j} \oplus W'_{2j}\ ,\ W''_{1j} \oplus W'_{3j}\ ,\ W''_{2j} \oplus W''_{3j} \quad \text{for every } j \in [3].$$

Referring back to Lemma 4.5.2, the Z_j here are separated linear and contain equal parts of an W_1W_2-, W_1W_3-, and W_2W_3-scheme (i.e., a pure symmetric two-way scheme). Since this caching scheme is symmetric it is enough to show there exists a response from the server of size at most $\frac{10}{6}F$ for each "demand type"[7].

When $\mathbf{d} = (i,i,i)$ for some $i \in [N]$, $X_{\mathbf{d}} = X_{iii} = W_i$ which is F bits.

When $\mathbf{d} = (i,j,k)$ for some $i,j,k \in [N]$ such that exactly two of the user requests are equal, there exists $X_{\mathbf{d}}$ of size $\frac{10}{6}F$ such that all users can reconstruct their requested files. For instance, take

$$X_{112} = (W'_{11} \oplus W'_{12}), W''_{11}, W''_{12}, W'_{13}, W''_{13}, W'_{21}, W'_{22}, W''_{21}, W''_{22}, W''_{23}.$$

[7] Referring to the number of unique files in the demand vector. For $N = K = 3$ there are 3 demand types, when all users demand the same file, when only two demand the same file, and when all demand distinct files.

Here X_{112} is comprised of 10 blocks each of size $F/6$ bits. Since $W'_{21} \subset X_{112}$ and $(W'_{11} \oplus W'_{21}) \subset Z_1$, then $W'_{11} \subset (X_{112},\, Z_1)$ and hence $W'_{12} \subset (X_{112},\, Z_1)$. Thus, $W_1 \subset (X_{112},\, Z_1)$. A similar argument shows that $W_1 \subset (X_{112},\, Z_2)$. Note that X_{112} contains all sections of W_2 other than W'_{23}. Since $W'_{13} \subset X_{112}$ and $(W'_{13} \oplus W'_{23}) \subset Z_3$, then $W'_{23} \subset (X_{112},\, Z_3)$. As a result $W_2 \subset (X_{112},\, Z_3)$.

When $\mathbf{d} = (i, j, k)$ for some distinct $i, j, k \in [N]$, there exists $X_{\mathbf{d}}$ of size $\frac{10}{6}F$ such that all users can reconstruct their requested files. For instance let $X_{123} = (A, B, C, D, E)$ where

$$\begin{aligned}
A &= W'_{12}, W''_{13}, W'_{21}, W''_{23}, W'_{31}, W''_{32}\,, \\
B &= (W''_{12} \oplus W''_{21} \oplus W''_{31}), \\
C &= (W'_{13} \oplus W''_{21} \oplus W''_{31}), \\
D &= (W''_{21} \oplus W''_{12} \oplus W'_{32}), \\
E &= (W''_{21} \oplus W''_{31} \oplus W''_{12} \oplus W'_{32} \oplus W'_{13} \oplus W'_{23}).
\end{aligned}$$

Here X_{112} is comprised of 10 blocks each of size $F/6$ bits. Then

$$\begin{aligned}
B \oplus E &= (W'_{32} \oplus W'_{13} \oplus W'_{23}), \\
C \oplus E &= (W'_{23} \oplus W''_{12} \oplus W'_{32}), \\
D \oplus E &= (W''_{31} \oplus W'_{13} \oplus W'_{23}).
\end{aligned}$$

We have that $(W'_{11}, W''_{11}, W'_{12}, W''_{13}) \subset (Z_1,\, A)$. Since $(W''_{21} \oplus W''_{31}) \subset Z_1$, then $W''_{12} \subset (Z_1,\, B)$ and $W'_{13} \subset (Z_1,\, C)$. Hence, $W_1 \subset (Z_1,\, X_{123})$. Similarly, we have $(W'_{22}, W''_{22}, W'_{21}, W''_{31}) \subset (Z_2,\, A)$. Since $(W''_{12} \oplus W'_{32}) \subset Z_2$, then $W'_{23} \subset (Z_2,\, D)$ and $W'_{23} \subset \big(Z_2,\, (C \oplus E)\big)$. Thus, $W_2 \subset (Z_2,\, X_{123})$. Lastly, we have $(W'_{33}, W''_{33}, W'_{31}, W''_{32}) \subset (Z_3,\, A)$. Since $(W'_{13} \oplus W'_{23}) \subset Z_3$, then $W'_{32} \subset \big(Z_3,\, (B \oplus E)\big)$ and $W''_{31} \subset \big(Z_3,\, (D \oplus E)\big)$. Consequently, $W_3 \subset (Z_3,\, X_{123})$ and all users can reconstruct their requested file with the given server response. □

Theorem 4.6.1 gives us an upper-bound on the memory-rate tradeoff for the $N = K = 3$ case. Given that $(M, R) = (1/3,\, 2)$ is achievable by [3], we know the line connecting $(M, R) = (1/3,\, 2)$ and $(M, R) = (1/2,\, 5/3)$, i.e, $6M + 3R = 8$, is achievable for $1/3 \leq M \leq 1/2$. Since, Tian proved the bound $6M + 3R \geq 8$ in

[9] for $N = K = 3$, the memory-rate tradeoff is completely characterized with the bound $6M + 3R \geq 8$ when $N = K = 3$ and $1/3 \leq M \leq 1/2$.

In [7], it is shown that $(M,R) = (1,\ 1)$ is achievable for $N = K = 3$, given the results of Theorem 4.6.1, it is likely that the other bound necessary to completely characterize the memory-rate tradeoff is $4M + 3R \geq 7$ for $1/2 \leq M \leq 1$. In Section 4.9 we show the closest bound we are able to derive with a linear separation assumption on the caches is $4M + 3R \geq 7 - \frac{1}{6}$.

4.7 A Discoordination Bound for $N = K = 3$

In this section we will prove a memory-rate bound for the linear coded caching problem in the $N = K = 3$ case. This bounds involves a discoordination term and can be considered an application of our work in Chapter 3 to the linear coded caching problem.

Theorem 4.7.1. *Consider the linear coded caching problem for $N = K = 3$ and where F is finite. Then*

$$2R + 3M \geq 5 - \frac{1}{F}\text{DisCoord}^{\text{avg}}(W_i, W_j, Z_k), \tag{4.7.1}$$

for $i, j, k \in [3]$ and $i \neq j$.

We organize this computation into a few lemmas.

Lemma 4.7.2. *Consider the linear coded caching problem for $N = K = 3$ and where F is finite. Then setting*

$$P_1 = (X_{123}, Z_1), \quad P_2 = (X_{213}, Z_2)$$

we have

$$2RF + 3MF \geq \dim(P_1) + \dim(P_2) + \dim(Z_3). \tag{4.7.2}$$

Proof. We have

$$2RF + 3MF \geq \dim(X_{123}) + \dim(Z_1) + \dim(X_{213}) + \dim(Z_2) + \dim(Z_3).$$

By the dimension formula

$$\dim(X_{123}) + \dim(Z_1) \geq \dim(X_{123} + Z_1) = \dim(P_1);$$

similarly

$$\dim(X_{213}) + \dim(Z_2) \geq \dim(P_2).$$

Combining the three equations displayed above yields the lemma. □

We remark that (4.7.2) would hold with equality if we add $\dim(X_{123} \cap Z_1)$ and $\dim(X_{213} \cap Z_2)$ to the right-hand-side.

Lemma 4.7.3. *Consider the hypothesis and notation of Lemma 4.7.2. Then*

$$2RF + 3MF \geq 4F + \dim^{\mathcal{U}/W_1}([P_1] \cap [P_2]) + \dim^{\mathcal{U}}\big((P_1 + P_2) \cap Z_3\big). \qquad (4.7.3)$$

Proof. By the dimension formula,

$$\dim(P_1) + \dim(P_2) = \dim(P_1 + P_2) + \dim(P_1 \cap P_2),$$

then the right-hand-side of (4.7.2) can be written as

$$\dim(P_1 + P_2 + Z_3) + \dim\big((P_1 + P_2) \cap Z_3\big) + \dim(P_1 \cap P_2).$$

But $P_1 + P_2 + Z_3$ implies X_{123}, Z_1, Z_2, Z_3 whose sum is all of $\mathcal{U}$. Hence

$$2RF + 3MF \geq 3F + \dim\big((P_1 + P_2) \cap Z_3\big) + \dim(P_1 \cap P_2).$$

Since P_1 and P_2 both imply W_1 we have

$$\dim(P_1 \cap P_2) = \dim^{\mathcal{U}/W_1}([P_1 \cap P_2]_{W_1}) + \dim(W_1) = \dim^{\mathcal{U}/W_1}([P_1] \cap [P_2]) + F,$$

and (4.7.3) follows. □

Next we study the first term on the right-hand-side of (4.7.3).

Lemma 4.7.4. *Consider the hypothesis and notation of Lemma 4.7.3. Then*

$$\dim^{\mathcal{U}/W_1}([P_1]\cap[P_2]) = \dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big) + t_1 + t_2 - \delta,$$

where

$$\delta = \mathrm{DisCoord}^{\mathcal{U}/W_1}([P_1],[P_2],[Z_3]) \tag{4.7.4}$$

and t_1, t_2 are the non-negative terms

$$t_1 = \dim^{\mathcal{U}/(W_1+W_3+Z_3)}\big([P_1]\cap[P_2]\big), \quad t_2 = \dim^{\mathcal{U}/W_1}([P_1]\cap[P_2]\cap[Z_3]).$$

In particular,

$$\dim^{\mathcal{U}/W_1}([P_1]\cap[P_2]) \geq \dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big) - \delta. \tag{4.7.5}$$

Proof. By Corollary 3.6.1 in the universe $\mathcal{U}/W_1$ and its three linear subspaces $[P_1],[P_2],[Z_3]$, we have

$$\begin{aligned}\dim^{\mathcal{U}/W_1}([P_1]\cap[P_2]) = &\dim^{\mathcal{U}/(W_1+Z_3)}\big([P_1]\cap[P_2]\big)\\ &+\dim^{\mathcal{U}/W_2}([P_1]\cap[P_2]\cap[Z_3])\\ &-\mathrm{DisCoord}^{\mathcal{U}/W_1}([P_1],[P_2],[Z_3]).\end{aligned}$$

Since both P_1 and P_2 contain W_3 when given Z_3 we have

$$[W_3]_{W_1+Z_3} \subset \big([P_1]_{W_1+Z_3}\cap[P_2]_{W_1+Z_3}\big),$$

hence

$$\dim^{\mathcal{U}/(W_1+Z_3)}\big([P_1]\cap[P_2]\big) = \dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big) + \dim^{\mathcal{U}/(W_3+W_1+Z_3)}\big([P_1]\cap[P_2]\big).$$

The equality in the lemma follows. □

Lemma 4.7.5. *Consider the hypothesis and notation of Lemma 4.7.2. Then*

$$2RF + 3MF \geq 5F + \dim^{\mathcal{U}/(W_1+W_2)}\big[(P_1+P_2)\cap Z_3]\big) + \dim(W_2\cap Z_3) + s_1 + s_2 - \delta, \tag{4.7.6}$$

where δ is the discoordination term given in (4.7.4)*, and*

$$s_1 = \dim(W_1+W_3+Z_3) - \dim(W_1+W_2+Z_3), \quad s_2 = \dim(W_1 \cap Z_3) - \dim(W_2 \cap Z_3). \tag{4.7.7}$$

Note that for the term $\dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big)$, we first calculate the intersection $(P_1+P_2)\cap Z_3$ in $\mathcal{U}$, and then consider the image of this intersection in $\mathcal{U}/(W_1+W_2)$. This distinction is important, for instance, in the optimal scheme for $M = 1/3$ (from [3]), the term equals 0, whereas the dimension of $[P_1+P_2]\cap[Z_3]$ in $\mathcal{U}/(W_1+W_2)$ equals $F/3$.

We also remark that s_1, s_2 in (4.7.7) cancel under symmetrization.

Proof of Lemma 4.7.5. Consider the second term on the right-hand-side of (4.7.3): since W_1, W_2 are both implied by P_1+P_2 (since P_1+P_2 contains X_{123}, Z_1, Z_2), we have

$$\dim\big((P_1+P_2)\cap Z_3\big) = \dim\big((W_1+W_2)\cap Z_3\big) + \dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big).$$

Combining this with Lemma 4.7.3 and Lemma 4.7.4, we have

$$\begin{aligned} 2RF+3MF \geq 4F &+ \dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big) - \delta \\ &+ \dim\big((W_1+W_2)\cap Z_3\big) + \dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big). \end{aligned} \tag{4.7.8}$$

Now we take two of the terms above and notice the following simplification (modulo s_1, which drops out upon symmetrization):

$$\dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big) = \dim(W_3+W_1+Z_3) - \dim(W_1+Z_3),$$

and

$$\dim\big((W_1+W_2)\cap Z_3\big) = \dim(W_1+W_2) + \dim(Z_3) - \dim(W_1+W_2+Z_3),$$

adding them we get

$$\dim^{\mathcal{U}/W_1+Z_3}(W_3)+\dim\big((W_1+W_2)\cap Z_3\big)=s_1-\dim(W_1+Z_3)\\ +\dim(W_1+W_2)+\dim(Z_3).$$

But here $\dim(Z_3)-\dim(W_1+Z_3)=\dim(W_1\cap Z_3)-\dim(W_1)$ and $\dim(W_1+W_2)=2F$, then

$$\dim^{\mathcal{U}/(W_1+Z_3)}\big([W_3]\big)+\dim\big((W_1+W_2)\cap Z_3\big)=F+s_1+\dim(W_1\cap Z_3)\\ =F+s_1+s_2+\dim(W_2\cap Z_3).$$

Applying the above equality to (4.7.8) yields (4.7.6). □

We can now prove the main result of this section.

Proof of Theorem 4.7.1. With notation as in Lemma 4.7.2, we will use (4.7.6) and a seemingly crude bound on the discoordination that uses the Lifting Lemma. First we recall the following equality from the proof of Lemma 4.7.5

$$\dim\big((P_1+P_2)\cap Z_3\big)=\dim\big((W_1+W_2)\cap Z_3\big)+\dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big).$$

Since W_1 is implied by P_1+P_2 we also have

$$\dim\big((P_1+P_2)\cap Z_3\big)=\dim(W_1\cap Z_3)+\dim^{\mathcal{U}/W_1}\big([(P_1+P_2)\cap Z_3]\big).$$

Hence,

$$\dim^{\mathcal{U}/W_1}\big([(P_1+P_2)\cap Z_3]\big)=\dim\big((W_1+W_2)\cap Z_3\big)-\dim(W_1\cap Z_3)\\ +\dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big).$$

By (3.6.1) and since $\dim(W_1\cap W_2)=0$ we have

$$\dim\big((W_1+W_2)\cap Z_3\big)-\dim(W_1\cap Z_3)=\mathrm{DisCoord}(W_1,W_2,Z_3)+\dim(W_2\cap Z_3).$$

Combining the above two bounds we get

$$\begin{aligned}\dim^{\mathcal{U}/W_1}\big([(P_1+P_2)\cap Z_3]\big) =& \text{DisCoord}(W_1,W_2,Z_3)+\dim(W_2\cap Z_3)\\ &+\dim^{\mathcal{U}/(W_1+W_2)}\big([(P_1+P_2)\cap Z_3]\big).\end{aligned} \tag{4.7.9}$$

Since $W_1 \subset (P_1 \cap P_2)$, by Theorem 3.1.8

$$\delta = \text{DisCoord}^{\mathcal{U}/W_1}([P_1],[P_2],[Z_3]) = \text{DisCoord}(P_1,P_2,Z_3).$$

By Theorem 3.1.4 and the Lifting Lemma (Lemma 3.4.2) we have

$$\text{DisCoord}(P_1,P_2,Z_3) = \dim^{\mathcal{U}/S_2}\big([P_1+P_2]\cap[Z_3]\big) = \dim^{\mathcal{U}/S_2}\big([(P_1+P_2)\cap Z_3]\big)$$

for $S_2 = \mathcal{S}_2(P_1,P_2,Z_3)$. Then $W_1 \subset (P_1 \cap P_2) \subset S_2$, and we get

$$\delta = \dim^{\mathcal{U}/S_2}\big([(P_1+P_2)\cap Z_3]\big) \leq \dim^{\mathcal{U}/W_1}\big([(P_1+P_2)\cap Z_3]\big).$$

Combining the above inequality with (4.7.9) we get

$$\dim^{\mathcal{U}/W_1+W_2}\big([(P_1+P_2)\cap Z_3]\big)+\dim(W_2\cap Z_3)-\delta \geq -\text{DisCoord}(W_1,W_2,Z_3),$$

which considering the notation and hypothesis of Lemma 4.7.5 leads to

$$2RF+3MF \geq 5F+s_1+s_2-\text{DisCoord}(W_1,W_2,Z_3). \tag{4.7.10}$$

After symmetrization s_1 and s_2 equal zero and we get (4.7.1). □

4.8 A Hybrid Rank Count and Tian's Method

By definition of the linear coded caching problem, $\mathcal{U}$ has a decomposition $W_1,\ldots,W_N$. Therefore, using Lemma 4.5.2 we know the structure of the caches in the linear coded caching problem with $N = K = 3$.

There is a stronger remark we can make about the structure of the caches if the Z's are separated linear under a symmetric caching scheme[8]. In this case, F is

[8]Take an arbitrary scheme and form the symmetrized scheme that is $K!N! = 36$ times as long

divisible by 3 and the files have a decomposition

$$W_1 = A_1A_2A_3, \quad W_2 = B_1B_2B_3, \quad W_3 = C_1C_2C_3, \tag{4.8.1}$$

such that A_i, B_i, C_i are each linear subspaces of dimension $F/3$ in $\mathcal{U}$ and Z_i is a linear subspace of $\mathcal{U}_i = \text{Span}(A_i, B_i, C_i)$ which trivially has a decomposition A_i, B_i, C_i. Using Lemma 4.5.2 we get the following remark.

Remark 4.8.1. *Consider the linear coded caching problem with $N = K = 3$ such that the Z_i are separated linear under a symmetric caching scheme, the files have a decomposition as in* (4.8.1)*, and $Z_i \subset \text{Span}(A_i, B_i, C_i)$. Then for each $i \in [3]$ there exist subspaces A_i^j, B_i^j, C_i^j indexed on integers $1 \leq j \leq 5$ such that*

1. *$A_i^1, \ldots, A_i^5$ are linearly independent subspaces that span A_i, as are $B_i^1, \ldots, B_i^5 \subset B_i$ and $C_i^1, \ldots, C_i^5 \subset C_i$ with $\text{Span}(B_i^1, \ldots, B_i^5) = B_i$ and $\text{Span}(C_i^1, \ldots, C_i^5) = C_i$;*

2. *for some integers $r_1, \ldots, r_5$ such that r_3 is even and*

$$r_1 + r_2 + r_3 + r_4 + r_5 = F/3, \tag{4.8.2}$$

   we have

$$\dim(A_i^j) = \dim(B_i^j) = \dim(C_i^j) = r_j,$$

   for each $j \in [5]$.

3. *Let $A_i^3 = A_i'^3 A_i''^3$ where $\dim(A_i'^3) = \dim(A_i''^3) = r_3/2$ and define a similar decomposition for B_i^3 and C_i^3, then Z_i is spanned by:*

   - *$A_i^1 + B_i^1 + C_i^1$ (an individual scheme of size $3r_1$ bits);*
   - *$A_i^2 \oplus B_i^2$, $B_i^2 \oplus C_i^2$ (a Tian scheme of size $2r_2$ bits);*
   - *$A_i'^3 \oplus B_i'^3$, $A_i''^3 \oplus C_i'^3$, $B_i''^3 \oplus C_i''^3$ (a symmetric two-way scheme of size $\frac{3}{2}r_3$ bits); and*
   - *$A_i^4 \oplus B_i^4 \oplus C_i^4$ (a triple scheme of size r_4 bits).*

(see Section 4.4.2).

4. for each $i \in [3]$,

$$\dim(Z_i) = MF = 3r_1 + 2r_2 + \frac{3}{2}r_3 + r_4. \tag{4.8.3}$$

Note that in Remark 4.8.1, A_i^5 is all the "leftovers" in A_i not spanned by $A_i^1, A_i^2, A_i^3, A_i^4$. A similar statement holds for B_i^5 and C_i^5. Since A_i^5, B_i^5, C_i^5 are disjoint from Z_i they can be considered as the parts of $\mathcal{U}_i$ "ignored" by Z_i.

Notice that if we choose a single scheme above, i.e., $r_j = F/3$ for only one value of $j \in [5]$, then

- If $r_1 = F/3$, i.e., we use a pure individual scheme, then $M = 1$ and $R = 1$ by [7]. This scheme is optimal by the bound $2M + 3R \geq 5$ shown in [9, 10].
- If $r_3 = F/3$, i.e., we use a pure symmetric two-way scheme, then $M = 1/2$ and $R = 5/3$ by Theorem 4.6.1. This scheme is optimal since $2R + 3M \geq 5 - 1/6$ by (4.8.4).
- If $r_4 = F/3$, i.e., we use a pure triple scheme, then $M = 1/3$ and $R = 2$ by [3]. This scheme is optimal by the trivial bound $3M + R \geq 3$.
- If $r_2 = F/3$, i.e., we use a pure Tian scheme, then $M = 2/3$ and (assuming the scheme is separated) Theorem 4.3.3 shows that $2R + 3M \geq 5$. This, $R \geq 3/2$, which is *worse* than a convex combination of $(1/2, 5/3)$ of $(1,1)$ which gives $(2/3, 13/9)$.

In other words, we know that three of the pure schemes above are optimal, and the pure Tian scheme cannot do as well as a convex combination of a pure individual scheme and a pure symmetric two-way scheme.

By applying Theorem 4.7.1 to the Z_i in the context of Remark 4.8.1 we are able to drive a new lower bound for the memory-rate tradeoff.

Corollary 4.8.2. *For the linear coded caching problem with $N = K = 3$ and separated linear caches, with integers r_2 and r_3 as in Remark 4.8.1,*

$$3MF + 2RF \geq 5F - r_2 - \frac{r_3}{2}. \tag{4.8.4}$$

Proof. It is enough to directly use (4.7.10) from the proof of Theorem 4.7.1, which gives

$$3MF + 2RF \geq 5F + \dim(A+C+Z_3) - \dim(A+B+Z_3) \\ + \dim(A \cap Z_3) - \dim(B \cap Z_3) - \text{DisCoord}(A,B,Z_3).$$

Since we are considering a symmetric caching scheme and by Corollary 3.1.7 the bound reduces to

$$3MF + 2RF \geq 5F - \dim\big((A+Z_3) \cap (B+Z_3)\big) + \dim(Z_3).$$

By Remark 4.8.1, $(A+Z_3) \cap (B+Z_3) \;=\; Z_3 + A_3^2 + A_3^{'3}$ and hence

$$3MF + 2RF \geq 5F - \dim(Z_3) - \dim(A_3^2) - \dim(A_3^{'3}) + \dim(Z_3).$$

Given that $\dim(A_3^2) = r_2$ and $\dim(A_3^{'3}) = \frac{1}{2}r_3$, (4.8.4) follows. □

Using a combination of Tian's argument from Theorem 4.3.3 and a matrix rank argument, we get a new memory-rate bound. This is the main computation of this Section.

Theorem 4.8.3. *For the linear coded caching problem with $N = K = 3$ and separated linear caches, with integers r_3 and r_5 as in Remark 4.8.1 we have,*

$$3MF + 2RF \geq 5F - \frac{3}{2}r_3 + 3r_5. \tag{4.8.5}$$

Note that, unlike (4.8.4), the bound above is not tight for $r_3 = F/3$ (i.e. a pure symmetric two-way scheme).

Our proof for Theorem 4.8.3 is similar to Tian's argument for Theorem 4.3.3. However, we need to do some "preprocessing" of the analog of the matrix G in our proof of Theorem 4.3.3 that was given in Section 4.3. Consider a pure triple scheme, meaning $r_4 = F/3$ and $r_j = 0$ for all other $j \in [5]$. Mimicking the proof of Theorem 4.3.3, we see that $X_{123} + Z_1$ contains the subspaces $A_1, A_2, A_3, (B_1 \oplus C_1)$,

and hence the matrices in (4.3.3) and (4.3.4) are at most of rank

$$(M+R'-4/3)F,$$

and conclude that

$$R' \leq 3(M+R'-4/3)F,$$

which gives the bound $2R'+3M \geq 4$. Furthermore, if $r_5 = F/3$, meaning each $Z_i = 0$, this type of argument would show

$$R' \leq 3(M+R'-3/3)F$$

and hence $2R'+3M \geq 3$. Similarly, for general $r_1,\ldots,r_5$, this method would show

$$R'F \leq 3(MF+R'F-F-2r_1-2r_2-\frac{3}{2}r_3-r_4),$$

which yields the lower bound

$$2R'F+3MF \geq 5F-\frac{3}{2}r_3-3r_4-6r_5.$$

A better approach will be a "hybrid" approach. First, we directly reason about the matrix G whose row space equals $\iota(X_{123})$, namely about its parts corresponding to the A_i^4, B_i^4, C_i^4 and the A_i^5, B_i^5, C_i^5 for $i \in [3]$. Then, we apply Tian's method to the remaining parts of G.

Proof of Theorem 4.8.3. Similar to the proof of Theorem 4.3.3, let us specify a basis for $W = W_1 + W_2 + W_3$. Consider a basis, $\mathcal{W}$, for W consisting of five parts based on the decomposition of the Z_i in Remark 4.8.1. Note that

$$W = \sum_{j=1}^{5} \Big(\sum_{i=1}^{3} (A_i^j + B_i^j + C_i^j) \Big).$$

For each $i \in [3]$ and $j = 1,2,5$, let $\mathcal{A}_i^j$ be a basis for A_i^j, and let $\mathcal{A}^j$ be the union of $\mathcal{A}_1^j, \mathcal{A}_2^j, \mathcal{A}_3^j$. Similarly for $\mathcal{B}$ or $\mathcal{C}$ replacing $\mathcal{A}$ and B or C replacing A everywhere.

Let $\mathcal{W}^j$ be the union of $\mathcal{A}^j, \mathcal{B}^j, \mathcal{C}^j$, so that we have

$$\mathcal{W}^j = \bigcup_{i\in[3]} \mathcal{A}_i^j \cup \mathcal{B}_i^j \cup \mathcal{C}_i^j \quad \text{for all } j = 1,2,5.$$

For each $i \in [3]$, let $\mathcal{A}_i'^3$ and $\mathcal{A}_i''^3$ be arbitrary bases of $A_i'^3$ and $A_i''^3$, respectively. Let $\mathcal{A}_i^3$ be the union of $\mathcal{A}_i'^3$ and $\mathcal{A}_i''^3$ and let $\mathcal{A}^3$ be the union of the $\mathcal{A}_i^3$. Similarly for $\mathcal{B}$ or $\mathcal{C}$ replacing $\mathcal{A}$ and B or C replacing A everywhere. Now, let $\mathcal{W}^3$ be the union of $\mathcal{A}^3, \mathcal{B}^3, \mathcal{C}^3$, so that we have

$$\mathcal{W}^3 = \bigcup_{i\in[3]} \mathcal{A}_i'^3 \cup \mathcal{A}_i''^3 \cup \mathcal{B}_i'^3 \cup \mathcal{B}_i''^3 \cup \mathcal{C}_i'^3 \cup \mathcal{C}_i''^3.$$

For $j = 4$ we take a different approach; for each $i \in [3]$, let $\mathcal{A}_i^4, \mathcal{B}_i^4$ respectively be arbitrary bases for A_i^4, B_i^4, and let $\tilde{\mathcal{C}}_i^4$ be an arbitrary basis for $A_i^4 \oplus B_i^4 \oplus C_i^4$ (recall the meaning of $A_i^4 \oplus B_i^4 \oplus C_i^4$ from Definition 4.5.1). Let $\mathcal{A}^4$ be the union of the $\mathcal{A}_i^4$, and similarly for $\mathcal{B}^4$ and $\tilde{\mathcal{C}}^4$. Let $\mathcal{W}^4$ be the union of these sets, then we have

$$\mathcal{W}^4 = \Big(\bigcup_{i\in[3]} \mathcal{A}_i^4 \cup \mathcal{B}_i^4 \Big) \quad \cup \quad \Big(\bigcup_{i\in[3]} \tilde{\mathcal{C}}_i^4 \Big).$$

Finally, let $\mathcal{W}$ be the union of the $\mathcal{W}^j$, which for block purposes we arrange in the order $\mathcal{W}^1, \ldots, \mathcal{W}^5$, so

$$\mathcal{W} \;=\; \mathcal{W}^1 \cup \mathcal{W}^2 \cup \mathcal{W}^3 \cup \mathcal{W}^4 \cup \mathcal{W}^5.$$

As in the proof of Theorem 4.3.3, this basis $\mathcal{W}$ of W gives an isomorphism

$$\iota = \iota_{\mathcal{W}} \colon W \to \mathbb{F}^{3F} \;\text{ with } \mathbb{F} = \mathbb{Z}/2\mathbb{Z}.$$

Note that for each $i \in [3]$, the vectors in $\iota(Z_i)$ have zeros in all their components corresponding to the basis elements in all $\mathcal{W}^5$, and of those in $\mathcal{A}^4$ and $\mathcal{B}^4$.

Let us describe a set of vectors in X_{123} that are linearly independent; their span will be a subspace of X_{123}, namely $\hat{X}_{123}$. Our goal is to describe vectors such that $\iota(\hat{X}_{123})$ has a convenient form which will be exploited and enables us to employ a hybrid strategy.

Consider any basis vector $a \in \mathcal{A}_1^5$. Since user 1 must be able to infer a from X_{123} and Z_1, we have $a = x + z$ where $x \in X_{123}$ and $z \in Z_1$. It follows that $-x = a - z$, and hence $\iota(-x)$—viewed as a block vector that breaks the basis $\mathcal{W}$ into its $\mathcal{W}^1, \ldots, \mathcal{W}^5$ blocks—is of the form:

$$\begin{bmatrix} \ell^1 & \ell^2 & \ell^3 & \ell^4 & e_r \end{bmatrix},$$

where e_r is one of the standard basis vectors in the $\mathcal{W}^5$ block (in particular, in the $\mathcal{A}_1^5$ part of $\mathcal{W}^5$), and where the $-\ell^j$ corresponds to the part of $\iota(z)$ in the $\mathcal{W}^j$ block of the basis $\mathcal{W}$; furthermore, as remarked before, ℓ^4 has zeros in the components corresponding to vectors in $\mathcal{A}^4$ and $\mathcal{B}^4$.

Doing this for each basis vector in $\mathcal{A}_1^5$, and similarly for the rest of the basis vectors in $\mathcal{W}^5$ we get a set of vectors in X_{123} whose image under ι, when arranged as row vectors, yields a block matrix of the form:

$$\begin{bmatrix} L^1 & L^2 & L^3 & L^4 & I \end{bmatrix}, \tag{4.8.6}$$

where I is a $9r_5F \times 9r_5F$ identity matrix, and for $j \in [4]$, L^j is a block matrix with $9r_5F$ rows. These rows are linearly independent because of the I in the block form above.

Now, note that user 1 can infer each element, $a \in \mathcal{A}_i^4$ with $i \in [3]$ from Z_1 and X_{123}. Therefore, $a = z + x$ with $z \in Z_1$ and $x \in X_{123}$, hence $\iota(x) = \iota(a) - \iota(z)$ gives us vectors of the form

$$\begin{bmatrix} \ell^1 & \ell^2 & \ell^3 & e_r + \ell^4 & 0_{\mathcal{W}^5} \end{bmatrix},$$

where e_r is the standard basis vector corresponding to $a \in \mathcal{A}^4$, $0_{\mathcal{W}^5}$ are the zeros for components corresponding to $\mathcal{W}^5$, and the ℓ^j result from $-\iota(z)$. Observe that ℓ^4 has zero components in the positions corresponding to $\mathcal{A}^4$ and $\mathcal{B}^4$, only possibly nonzero in those components corresponding to $\tilde{\mathcal{C}}^4$. Doing the same for all $b \in \mathcal{B}_i^4$ with $i \in [3]$, we get elements of X_{123} such that ι of these elements, arranged as row vectors, is of the form

$$\left[\begin{array}{ccc|c|c|c} P^1 & P^2 & P^3 & Q' & I & 0 \end{array}\right] \tag{4.8.7}$$

where 0 is a zero matrix of size $6r_4 \times 9r_5$ columns and I in the block an identity matrix of size $6r_4 \times 6r_4$, and Q' is the matrix of components corresponding to elements of $\tilde{\mathcal{C}}^4$.

All the rows of the matrices in (4.8.7) and (4.8.6) are linearly independent with the following argument: when we combine these matrices we get a matrix

$$\left[\begin{array}{ccc|cc|c} P^1 & P^2 & P^3 & Q' & I & 0 \\ L^1 & L^2 & L^3 & L_4 & & I \end{array}\right]. \tag{4.8.8}$$

L^4 is equivalent to $[Q \quad 0]$ where 0 is a zero matrix of size $9r_5 \times 6r_4$ and Q is the matrix of components corresponding to elements of $\tilde{\mathcal{C}}^4$, since the Z_i have zero components corresponding to elements of $\mathcal{A}^4$ and $\mathcal{B}^4$. Hence we get a block matrix

$$\begin{bmatrix} P^1 & P^2 & P^3 & Q' & I & 0 \\ L^1 & L^2 & L^3 & Q & 0 & I \end{bmatrix}, \tag{4.8.9}$$

whose two right-most columns give a $(6r_4+9r_5) \times (6r_4+9r_5)$ identity matrix.

At this point we have identified a subspace X'_{123} of X_{123}, and a basis of X'_{123}, whose image under ι, viewed as row vectors, equals the block matrix in (4.8.9). Now list all of the vectors in $\mathcal{W}^1 \cup \mathcal{W}^2 \cup \mathcal{W}^3$ as a sequence in any order

$$v_1, \ldots, v_m$$

(note that here the subscripts do not refer to a scheme or a user). Each $v_k = x_k + z_k$ for some $x_k \in X_{123}$ and $z_k \in Z_1 \cup Z_2 \cup Z_3$. Let

$$\hat{X}_{123} = X'_{123} + \text{Span}(x_1, \ldots, x_m).$$

Now we create a matrix whose rowspace equals $\iota(\hat{X}_{123})$ as follows: we begin with the matrix in (4.8.9) and for $k \in [m]$ we add a row for each x_k such that

$$x_k \notin X'_{123} + \text{Span}(x_1, \ldots, x_{k-1})$$

using the same idea as before. Since $-x_k = v_k - z_k$, we add the row $\iota(-x_k) =$

$\iota(v_k) - \iota(z_k)$ which has the form

$$\begin{bmatrix} \ell^1 & \ell^2 & \ell^3 & \ell^4 & 0_{\mathcal{A}^4,\mathcal{B}^4} & 0_{\mathcal{W}^5} \end{bmatrix},$$

where ℓ^4 corresponds to the $\tilde{\mathcal{C}}^4$ part, and the subscripts on the 0's indicate the zeros corresponding to the $\mathcal{A}^4, \mathcal{B}^4$ and $\mathcal{W}^5$ components. Adding all such vectors x_k to obtain $\hat{X}_{123}$ we have that $\iota(\hat{X}_{123})$, viewed as row vectors, is the rowspace of a matrix

$$G = \begin{bmatrix} G''^1 & G''^2 & G''^3 & Q'' & 0 & 0 \\ P^1 & P^2 & P^3 & Q' & I & 0 \\ L^1 & L^2 & L^3 & Q & 0 & I \end{bmatrix}. \tag{4.8.10}$$

Setting

$$G'' = \begin{bmatrix} G''^1 & G''^2 & G''^3 \end{bmatrix},$$

we have

$$FR \geq \dim(X_{123}) \geq \dim(\hat{X}_{123}) = \mathrm{Rank}(G),$$

where

$$\mathrm{Rank}(G) = \mathrm{Rank}([G'' \quad Q'']) + 6r_4 + 9r_5 \geq \mathrm{Rank}(G'') + 6r_4 + 9r_5,$$

and hence

$$RF \geq R''F + 6r_4 + 9r_5, \quad \text{where} \quad R''F = \mathrm{Rank}(G''). \tag{4.8.11}$$

Our aim is to apply Tian's argument from Theorem 4.3.3 to G''. To do so, first we claim that

$$\iota\Big(\mathrm{Span}(x_1, \ldots, x_m)\Big)$$

lies entirely in the rowspace of $[G'' \quad Q'' \quad 0 \quad 0]$. Each x_k with

$$x_k \notin X'_{123} + \mathrm{Span}(x_1, \ldots, x_{k-1})$$

has $\iota(x_k)$ as one of the rows of $[G'' \quad Q'' \quad 0 \quad 0]$, by our construction above. However, if

$$x_k \in X'_{123} + \mathrm{Span}(x_1, \ldots, x_{k-1})$$

then $\iota(x_k)$ lies in some combination of the rowspace of G in (4.8.10). But since $x_k = v_k - z_k$, then x_k has zero components in positions corresponding to $\mathcal{A}^4, \mathcal{B}^4$ and $\mathcal{W}^5$; but since the two last columns of G are

$$\begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix},$$

and v_k corresponds to a vector in one of $\mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^3$, we have that $\iota(x_k) = \iota(v_k) - \iota(z_k)$ has zero component in the positions corresponding to $\mathcal{A}^4, \mathcal{B}^4$ and $\mathcal{W}^5$. Hence $\iota(x_k)$, which is a linear combination of rows in G, cannot involve the bottom two rows blocks, which correspond to $\iota(X'_{123})$.

Now we know

$$\iota\Big(\text{Span}(x_1, \ldots, x_m)\Big) = \text{Rowspace}\Big([G'' \quad Q'' \quad 0 \quad 0]\Big).$$

Consider the special case where all the vectors in $\tilde{\mathcal{C}}^4 = 0$, i.e., $A_i^4 \oplus B_i^4 \oplus C_i^4 = 0$ and $Z_i = Z_i'' = Z_i \cap \text{Span}(\mathcal{W}^1 \cup \mathcal{W}^2 \cup \mathcal{W}^3)$ for all $i \in [3]$. In this special case, user 1 can reconstruct A_i^j for all $j, i \in [3]$ and we may set $Q'' = 0$. Since each vector in $\mathcal{A}_i^j$ occurs in the sequence $v_1, \ldots, v_m$, we may compute the same values of $v_1, \ldots, v_m$ with $Q'' = 0$; we can replace $x_1, \ldots, x_m$ with the vectors $x_1'', \ldots, x_m''$ obtained by discarding the $\tilde{\mathcal{C}}^4$ components of $x_1, \ldots, x_m$ and still get a scheme that allows users to reconstruct the $\mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^3$ parts of their files; where the total memory user i needs to do this is

$$\dim(Z_i'') = M''F = MF - r_4.$$

Additionally, the dimension of

$$X_{123}'' = \text{Span}(x_1'', \ldots, x_m''),$$

which equals $R''F = \text{Rank}(G'')$.

Now we apply the same argument as in Theorem 4.3.3, to conclude that the

span of the columns of G'' corresponding to $\mathcal{B}_2^j, \mathcal{B}_3^j, \mathcal{C}_2^j, \mathcal{C}_3^j$ for all $j \in [3]$ is at most

$$M''F + R''F - 5r_1 - 5r_2 - \frac{9}{2}r_3.$$

Applying the same argument for Z_2 and Z_3 replacing Z_1, we conclude

$$R''F \leq 3(M''F + R''F - 5r_1 - 5r_2 - \frac{9}{2}r_3),$$

and hence

$$3M''F + 2R''F \geq 15r_1 + 15r_2 + \frac{27}{2}r_3.$$

Using (4.8.11) and the fact that $MF = M''F + r_4$ we have

$$\begin{aligned} 3MF + 2RF &\geq 3(M''F + r_4) + 2(R''F + 6r_4 + 9r_5) \\ &\geq 15r_1 + 15r_2 + \frac{27}{2}r_3 + 15r_4 + 18r_5 \\ &\geq 15(r_1 + r_2 + r_3 + r_4 + r_5) - \frac{3}{2}r_3 + 3r_5. \end{aligned}$$

Since $\sum_{i=1}^{5} r_i = F/3$, we conclude

$$3MF + 2RF \geq 5F - \frac{3}{2}r_3 + 3r_5.$$

□

4.9 Better Bounds and a Few Conjectures

4.9.1 Bounds

We can drive new bounds for the linear coded caching problem with separated caches using Remark 4.8.1, Corollary 4.8.2, and Theorem 4.8.3.

Corollary 4.9.1. *In the linear coded caching problem with $N = K = 3$ and separated linear caches we have*

$$6M + 5R \geq 11.$$

Proof. To obtain this corollary we add (4.8.5) to $3/2$ times (4.8.4) and add the

result to $3/2$ times (4.8.3) plus $-9/2$ times (4.8.2) which yields

$$5RF + 6MF \geq 11F + 3r_4 + 5r_5.$$

□

Similar to how the bound in this corollary was obtained, with a different combination of (4.8.5), (4.8.4), (4.8.3), and (4.8.2) we can also show

$$2R + 3M \geq 5 - \frac{1}{4}, \text{ and } \quad 3R + 4M \geq 7 - \frac{1}{6}.$$

With the linearity and separation assumptions we improve Tian's bounds (which were $2R + 3M \geq 5 - \frac{1}{3}$ and $3R + 4M \geq 7 - \frac{1}{3}$) but we are unable to show $3R + 4M \geq 7$.

The bound in Corollary 4.9.1 gives a slight improvement to Tian's bound $R + M \geq 2$, as both pass through the achievable point $(M, R) = (1, 1)$; however, the bound in Corollary 4.9.1 assumes that the caching scheme is separated linear; by contrast, Tian's bound is valid for any scheme, including non-linear schemes.

We remark that Tian's bound $R + M \geq 2$ has a short proof. In the context of the linear coded caching problem, $R + M \geq 2$ follows from the fact that

$$(2R + 2M)F \geq \dim(X_{123} + Z_1) + \dim(X_{213} + Z_2),$$

which by the dimension formula equals

$$\dim(X_{123} + Z_1 + X_{213} + Z_2) + \dim\big((X_{123} + Z_1) \cap (X_{213} + Z_2)\big),$$

where the first dimension equals $3F$, and the second dimension is at least that of W_1, namely F. For non-linear schemes this proof still holds, since the above lower bound on $2R + 2M$ becomes

$$H(X_{123}, Z_1, X_{213}, Z_2) + I\big((X_{123} + Z_1); (X_{213} + Z_2)\big),$$

which is bounded below by $3F + F$, using the fact that the two-way mutual information, $I(X;Y)$, of random variables X and Y is bounded from below by $H(Z)$ for

any Z that is implied by both X and Y.

4.9.2 Conjectures

We make the following conjectures:

1. One can improve the bound in Theorem 4.8.3 to

$$2RF + 3MF \geq 5F - (1/2)r_3 + 3r_6,$$

   which would be tight for $(M,R) = (1/2,\ 5/3)$, and would then imply that no separated linear scheme can improve upon a convex combination of the scheme achieving $(M,R) = (1/2,5/3)$ and the scheme achieving $(M,R) = (1,1)$.

2. Any optimal linear scheme is separated.

3. Recall, the line connecting $(M,R) = (1/2,5/3)$ and $(M,R) = (1,1)$ is $4M + 3R = 7$; we conjecture than

$$4M + 3R \geq 7$$

   holds for all $1/2 \leq M \leq 1$ under any linear caching scheme.

Our difficulty in attacking either conjecture (1) or (2) is the possible ways in which the X_{ijk} can involve XORs of the bits in A_i^j, B_i^j, C_i^j over different values of $j \in [4]$.

If conjecture (3) is shown, then $4M + 3R \geq 7$ holds for all schemes unless there is a non-linear scheme which improves upon this (we do not particularly conjecture one way or another on the existence of such a non-linear scheme).

Chapter 5

Conclusion

5.1 Conclusion

In this thesis, we begin by formalizing linear information theory; what are the linearity assumptions that allow us to represent the entropy of a random variable as the dimension of a linear subspace in some universe? In other words, what is a linear random variable? We discuss the notion of "coordination of linear subspaces" and define "discoordination" as a measure of the extent to which linear subspaces fail to be coordinated. After that, we generalize the proof of the dimension formula to quasi-increasing sequences of linear subspaces. Next, building upon our coordination results, we derive a closed-form expression for the discoordination of a collection of subspaces. Exploiting the peculiar properties of coordinating three linear subspaces, we show different (more useful) equalities involving the discoordination of three subspaces.

Then we move on to the coded caching problem. After formally defining the problem, we review the relevant literature and related results. We define linear coded caching and restate one of Tian's ideas from [9] as Theorem 4.3.3; we extend this idea to get a different lower bound for the memory-rate tradeoff. Furthermore, we completely characterize the memory rate tradeoff for $N = K = 3$ and $1/3 \leq M \leq 1/2$ by defining a new caching scheme that achieves the memory rate pair $(M,\, R) = (1/2, 5/3)$. Using our results in linear information theory, we derive a lower bound for the memory-rate tradeoff involving a discoordination term.

Combing our bounds we show $5R + 6M \geq 11$ for the linear coded caching problem with $N = K = 3$; this bound is an improvement on Tian's bound of $R + M \geq 2$, however our bound assumes the random variables in the coded caching problem are linear. All the new lower bounds on the memory-rate tradeoff in this thesis were, in effect, the result of a linear-algebraic approach to the coded caching problem.

5.2 Future Work

It is intriguing to know if one can define concepts analogous to discoordination for non-linear random variables and obtain our results without the linearity assumption. Coordinating four or more linear subspaces is much more complicated (see remark after Theorem 3.3.10) and there might not even be a decomposition result similar to that of Theorem 3.1.4. Therefore, finding a relation between $\text{DisCoord}(U_1, \ldots, U_m)$, $\dim(U_1 \cap \ldots \cap U_m)$, and $I(U_1; \ldots; U_m)$ for $m \geq 4$ is of great interest.

We conjectured that $4M + 3R \geq 7$ will completely characterize the memory-rate tradeoff in the linear coded caching problem for the $N = K = 3$ case and $1/2 \leq M \leq 1$. Needless to say, our (possible) continuation of this work should address this conjecture. Naturally, all the conjectures in Subsection 4.9.2 are good places to extend this work.

Our work demonstrated the application of linear information theory to the coded caching problem; another direction for future work can involve finding other applications of linear information theory.

5.3 Final Remarks

In writing this thesis, I read many others. Justifiably, one common trend in all of them was a lack of personal context. A master's thesis in computer science is likely a survey of research done over one or two years and reads more like an information dump rather than a story. I thought it's worth writing a couple of words in an attempt to give the story behind this work, so there is some idea of how this work came together and what the author thought of it.

Joel Friedman introduced me to the coded caching problem in late September

2020 after he attended the dissertation defense of Ali Saberali[1] as an examiner. Joel saw that characterizing the memory-rate tradeoff was an open problem, even for the small case with three users and three files. For the next couple of months, we reviewed the literature and attempted numerous unsuccessful approaches to tackling the problem. In early 2021, Joel came up with "discoordination", and by March we had more or less shown the theorems in Chapter 3 of this work. During the summer, I wrote the argument which appears in Section 4.7. Initially, there was a mistake in the argument which we didn't catch and led us to believe that we proved $3M + 2R \geq 5$. We found the mistake in yearly 2022 and fixed the argument. This led to all the sections in Chapter 4 related to Tian's work in [9].

My experience during my time as a graduate student was somewhat similar to that of Aegeus, following his son's (Theseus) departure to Crete. The imagery was eloquently put in [5] by Stephen Fry, where it's said: Aegeus had stood patiently every day on the cliffs overlooking the sea bearing his name, waiting for a sight of his son's ship. I don't relate my research experience to the tragic end of Aegeus, but his routine. Research can be a lot like standing on a cliff looking at the horizon for a ship. But with the promise that you return to stand on the same cliff the next day and all days after that to look upon the horizon until you spot the ship you want (the one with the white sails). This work is the result of patience as much as it is the result of effort. It's the result of persistence and trust as it's the result of knowledge and guidance. I'm happy with the ship I saw at the end of my routine, and I thank you, the reader, for being a part of it.

[1] Incidentally, one of the examiners of this thesis, Sathish Gopalakrishnan, was also an examiner of Saberali's dissertation (see [8]). In one of our meetings, Sathish kindly shared with me a very intriguing problem related to coded caching inspired by Saberali's work. We noted how he and Joel came up completely different questions (!) after examining the same work.

Bibliography

[1] M. F. Atiyah and I. G. MacDonald. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969. ISBN 978-0-201-40751-8. → page 17

[2] S. Axler. *Linear algebra done right*. Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015. ISBN 978-3-319-11079-0; 978-3-319-11080-6. doi:10.1007/978-3-319-11080-6. URL https://doi.org/10.1007/978-3-319-11080-6. → pages 2, 7, 10

[3] Z. Chen, P. Fan, and K. Letaief. Fundamental limits of caching: Improved bounds for users with small buffers. *IET Communications*, 10, 07 2016. doi:10.1049/iet-com.2015.1205. → pages 75, 76, 77, 95, 99, 103

[4] T. M. Cover and J. A. Thomas. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, July 2006. ISBN 0471241954. → page 17

[5] S. Fry. *Heroes: The myths of the Ancient Greek heroes retold*. Stephen Fry's Greek Myths. Penguin Books Limited, 2018. ISBN 9781405940382. URL https://books.google.ca/books?id=4QRuDwAAQBAJ. → page 116

[6] K. Jänich. *Linear algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1994. ISBN 0-387-94128-2. doi:10.1007/978-1-4612-4298-7. URL https://doi.org/10.1007/978-1-4612-4298-7. → pages 2, 3, 7, 10, 15, 23

[7] M. A. Maddah-Ali and U. Niesen. Fundamental limits of caching. *IEEE Transactions on Information Theory*, 60(5):2856–2867, 2014. doi:10.1109/TIT.2014.2306938. → pages 70, 71, 73, 74, 76, 77, 96, 103

[8] S. A. Saberali. *Coded caching: convex optimization and graph theoretical perspectives*. PhD thesis, University of British Columbia, 2020. URL https://open.library.ubc.ca/collections/ubctheses/24/items/1.0394742. → pages 70, 116

[9] C. Tian. Symmetry, outer bounds, and code constructions: A computer-aided investigation on the fundamental limits of caching. *Entropy*, 20(8), 2018. ISSN 1099-4300. doi:10.3390/e20080603. URL https://www.mdpi.com/1099-4300/20/8/603. → pages 6, 70, 75, 76, 77, 78, 82, 83, 84, 85, 96, 103, 114, 116

[10] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr. Characterizing the rate-memory tradeoff in cache networks within a factor of 2. *IEEE Trans. Inf. Theor.*, 65(1):647–663, Jan. 2019. ISSN 0018-9448. doi:10.1109/TIT.2018.2870566. URL https://doi.org/10.1109/TIT.2018.2870566. → pages 70, 74, 75, 76, 83, 103