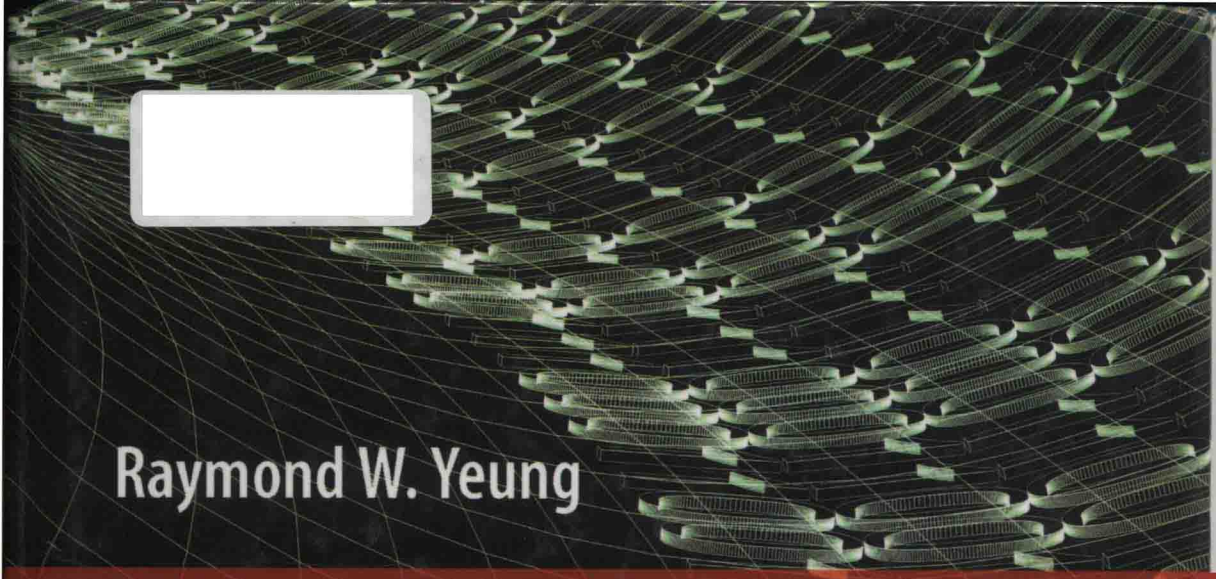


Raymond W. Yeung

INFORMATION THEORY AND NETWORK CODING

Springer

Information Theory and Network Coding

By Raymond W. Yeung

Information Theory and Network Coding consists of two parts: Components of Information Theory and Fundamentals of Network Coding Theory. Part I is a rigorous treatment of information theory for discrete and continuous systems. In addition to the classical topics, there are such modern topics as the I-Measure, Shannon-type and non-Shannon-type information inequalities, and a fundamental relation between entropy and group theory. With information theory as the foundation, Part II is a comprehensive treatment of network coding theory with detailed discussions on linear network codes, convolutional network codes, and multi-source network coding.

Other important features include:

- Derivations that are from the first principle
- A large number of examples throughout the book
- Many original exercise problems
- Easy-to-use chapter summaries
- Two parts that can be used separately or together for a comprehensive course

Information Theory and Network Coding is for senior undergraduate and graduate students in electrical engineering, computer science, and applied mathematics. This work can also be used as a reference for professional engineers in the area of communications.

ISBN 978-0-387-79233-0

9 780387 792330

springer.com

Yeung

INFORMATION THEORY AND NETWORK CODING

Raymond W. Yeung

Information Theory and Network Coding

Raymond W. Yeung
The Chinese University of Hong Kong
Department of Information Engineering
Hong Kong
People's Republic of China
whyeung@ie.cuhk.edu.hk

ISBN: 978-0-387-79233-0 e-ISBN: 978-0-387-79234-7

Library of Congress Control Number: 2008924472

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Information Theory and Network Coding

Information Technology: Transmission, Processing, and Storage

---

Information Theory and Network Coding
Raymond W. Yeung

Networks and Grids: Technology and Theory
Thomas G. Robertazzi

CDMA Radio with Repeaters
Joseph Shapira and Shmuel Miller

Digital Satellite Communications
Giovanni Corazza, ed.

Immersive Audio Signal Processing
Sunil Bharitkar and Chris Kyriakakis

Digital Signal Processing for Measurement Systems: Theory and Applications
Gabriele D'Antona and Alessandro Ferrero

Coding for Wireless Channels
Ezio Biglieri

Wireless Networks: Multiuser Detection in Cross-Layer Design
Christina Comaniciu, Narayan B. Mandayam and H. Vincent Poor

The Multimedia Internet
Stephen Weinstein

MIMO Signals and Systems
Horst J. Bessai

Multi-Carrier Digital Communications:
Theory and Applications of OFDM, 2nd Ed
Ahmad R.S. Bahai, Burton R. Saltzberg and Mustafa Ergen

Performance Analysis and Modeling of Digital Transmission Systems
William Turin

Wireless Communications Systems and Networks
Mohsen Guizani

Interference Avoidance Methods for Wireless Systems
Dimitrie C. Popescu and Christopher Rose

Stochastic Image Processing
Chee Sun Won and Robert M. Gray

Coded Modulation Systems
John B. Anderson and Arne Svensson

Communication System Design Using DSP Algorithms:
With Laboratory Experiments for the TMS320C6701 and TMS320C6711
Steven A. Tretter

(continued after index)

To my parents and my family

Preface

This book is an evolution from my book *A First Course in Information Theory* published in 2002 when network coding was still at its infancy. The last few years have witnessed the rapid development of network coding into a research field of its own in information science. With its root in information theory, network coding has not only brought about a paradigm shift in network communications at large, but also had significant influence on such specific research fields as coding theory, networking, switching, wireless communications, distributed data storage, cryptography, and optimization theory. While new applications of network coding keep emerging, the fundamental results that lay the foundation of the subject are more or less mature. One of the main goals of this book therefore is to present these results in a unifying and coherent manner.

While the previous book focused only on information theory for discrete random variables, the current book contains two new chapters on information theory for continuous random variables, namely the chapter on differential entropy and the chapter on continuous-valued channels. With these topics included, the book becomes more comprehensive and is more suitable to be used as a textbook for a course in an electrical engineering department.

What Is in This book

Out of the 21 chapters in this book, the first 16 chapters belong to Part I, *Components of Information Theory*, and the last 5 chapters belong to Part II, *Fundamentals of Network Coding*. Part I covers the basic topics in information theory and prepares the reader for the discussions in Part II. A brief rundown of the chapters will give a better idea of what is in this book.

Chapter 1 contains a high-level introduction to the contents of this book. First, there is a discussion on the nature of information theory and the main results in Shannon's original paper in 1948 which founded the field. There are also pointers to Shannon's biographies and his works.

Chapter 2 introduces Shannon's information measures for discrete random variables and their basic properties. Useful identities and inequalities in

information theory are derived and explained. Extra care is taken in handling joint distributions with zero probability masses. There is a section devoted to the discussion of maximum entropy distributions. The chapter ends with a section on the entropy rate of a stationary information source.

Chapter 3 is an introduction to the theory of I-Measure which establishes a one-to-one correspondence between Shannon's information measures and set theory. A number of examples are given to show how the use of information diagrams can simplify the proofs of many results in information theory. Such diagrams are becoming standard tools for solving information theory problems.

Chapter 4 is a discussion of zero-error data compression by uniquely decodable codes, with prefix codes as a special case. A proof of the entropy bound for prefix codes which involves neither the Kraft inequality nor the fundamental inequality is given. This proof facilitates the discussion of the redundancy of prefix codes.

Chapter 5 is a thorough treatment of weak typicality. The weak asymptotic equipartition property and the source coding theorem are discussed. An explanation of the fact that a good data compression scheme produces almost i.i.d. bits is given. There is also an introductory discussion of the Shannon–McMillan–Breiman theorem. The concept of weak typicality will be further developed in Chapter 10 for continuous random variables.

Chapter 6 contains a detailed discussion of strong typicality which applies to random variables with finite alphabets. The results developed in this chapter will be used for proving the channel coding theorem and the rate-distortion theorem in the next two chapters.

The discussion in Chapter 7 of the discrete memoryless channel is an enhancement of the discussion in the previous book. In particular, the new definition of the discrete memoryless channel enables rigorous formulation and analysis of coding schemes for such channels with or without feedback. The proof of the channel coding theorem uses a graphical model approach that helps explain the conditional independence of the random variables.

Chapter 8 is an introduction to rate-distortion theory. The version of the rate-distortion theorem here, proved by using strong typicality, is a stronger version of the original theorem obtained by Shannon.

In Chapter 9, the Blahut–Arimoto algorithms for computing the channel capacity and the rate-distortion function are discussed, and a simplified proof for convergence is given. Great care is taken in handling distributions with zero probability masses.

Chapters 10 and 11 are devoted to the discussion of information theory for continuous random variables. Chapter 10 introduces differential entropy and related information measures, and their basic properties are discussed. The asymptotic equipartition property for continuous random variables is proved. The last section on maximum differential entropy distributions echoes the section in Chapter 2 on maximum entropy distributions.

Chapter 11 discusses a variety of continuous-valued channels, with the continuous memoryless channel being the basic building block. In proving the capacity of the memoryless Gaussian channel, a careful justification is given for the existence of the differential entropy of the output random variable. Based on this result, the capacity of a system of parallel/correlated Gaussian channels is obtained. Heuristic arguments leading to the formula for the capacity of the bandlimited white/colored Gaussian channel are given. The chapter ends with a proof of the fact that zero-mean Gaussian noise is the worst additive noise.

Chapter 12 explores the structure of the I-Measure for Markov structures. Set-theoretic characterizations of full conditional independence and Markov random field are discussed. The treatment of Markov random field here maybe too specialized for the average reader, but the structure of the I-Measure and the simplicity of the information diagram for a Markov chain are best explained as a special case of a Markov random field.

Information inequalities are sometimes called the laws of information theory because they govern the impossibilities in information theory. In Chapter 13, the geometrical meaning of information inequalities and the relation between information inequalities and conditional independence are explained in depth. The framework for information inequalities discussed here is the basis of the next two chapters.

Chapter 14 explains how the problem of proving information inequalities can be formulated as a linear programming problem. This leads to a complete characterization of all information inequalities provable by conventional techniques. These inequalities, called Shannon-type inequalities, can be proved by the World Wide Web available software package ITIP. It is also shown how Shannon-type inequalities can be used to tackle the implication problem of conditional independence in probability theory.

Shannon-type inequalities are all the information inequalities known during the first half century of information theory. In the late 1990s, a few new inequalities, called non-Shannon-type inequalities, were discovered. These inequalities imply the existence of laws in information theory beyond those laid down by Shannon. In Chapter 15, we discuss these inequalities and their applications.

Chapter 16 explains an intriguing relation between information theory and group theory. Specifically, for every information inequality satisfied by any joint probability distribution, there is a corresponding group inequality satisfied by any finite group and its subgroups and vice versa. Inequalities of the latter type govern the orders of any finite group and their subgroups. Group-theoretic proofs of Shannon-type information inequalities are given. At the end of the chapter, a group inequality is obtained from a non-Shannon-type inequality discussed in Chapter 15. The meaning and the implication of this inequality are yet to be understood.

Chapter 17 starts Part II of the book with a discussion of the butterfly network, the primary example in network coding. Variations of the butterfly

network are analyzed in detail. The advantage of network coding over store-and-forward in wireless and satellite communications is explained through a simple example. We also explain why network coding with multiple information sources is substantially different from network coding with a single information source.

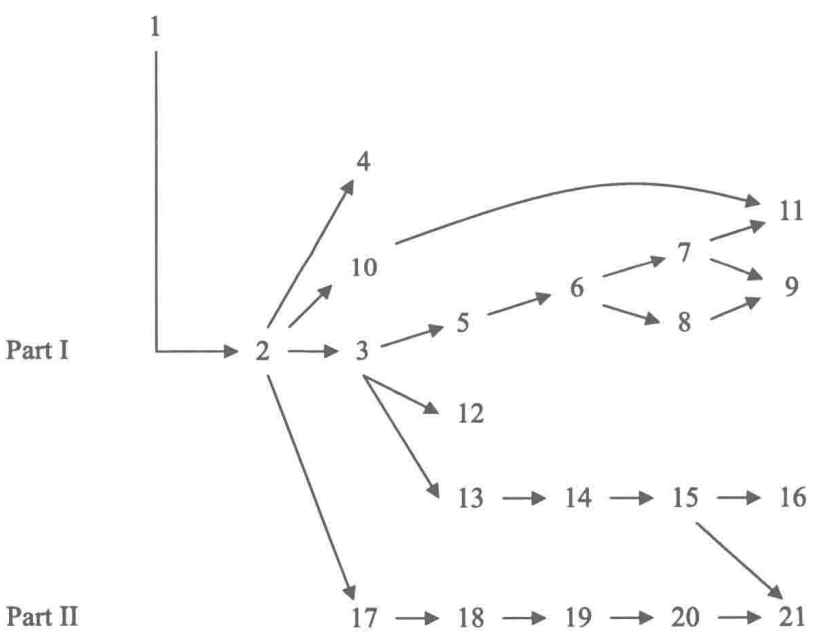
In Chapter 18, the fundamental bound for single-source network coding, called the max-flow bound, is explained in detail. The bound is established for a general class of network codes.

In Chapter 19, we discuss various classes of linear network codes on acyclic networks that achieve the max-flow bound to different extents. Static network codes, a special class of linear network codes that achieves the max-flow bound in the presence of channel failure, are also discussed. Polynomial-time algorithms for constructing these codes are presented.

In Chapter 20, we formulate and analyze convolutional network codes on cyclic networks. The existence of such codes that achieve the max-flow bound is proved.

Network coding theory is further developed in Chapter 21. The scenario when more than one information source are multicast in a point-to-point acyclic network is discussed. An implicit characterization of the achievable information rate region which involves the framework for information inequalities developed in Part I is proved.

How to Use This book



Part I of this book by itself may be regarded as a comprehensive textbook in information theory. The main reason why the book is in the present form is because in my opinion, the discussion of network coding in Part II is incomplete without Part I. Nevertheless, except for Chapter 21 on multi-source network coding, Part II by itself may be used satisfactorily as a textbook on single-source network coding.

An elementary course on probability theory and an elementary course on linear algebra are prerequisites to Part I and Part II, respectively. For Chapter 11, some background knowledge on digital communication systems would be helpful, and for Chapter 20, some prior exposure to discrete-time linear systems is necessary. The reader is recommended to read the chapters according to the above chart. However, one will not have too much difficulty jumping around in the book because there should be sufficient references to the previous relevant sections.

This book inherits the writing style from the previous book, namely that all the derivations are from the first principle. The book contains a large number of examples, where important points are very often made. To facilitate the use of the book, there is a summary at the end of each chapter.

This book can be used as a textbook or a reference book. As a textbook, it is ideal for a two-semester course, with the first and second semesters covering selected topics from Part I and Part II, respectively. A comprehensive instructor's manual is available upon request. Please contact the author at whyeung@ie.cuhk.edu.hk for information and access.

Just like any other lengthy document, this book for sure contains errors and omissions. To alleviate the problem, an errata will be maintained at the book homepage http://www.ie.cuhk.edu.hk/IT_book2/.

Hong Kong, China
December, 2007

Raymond W. Yeung

Acknowledgments

The current book, an expansion of my previous book *A First Course in Information Theory*, was written within the year 2007. Thanks to the generous support of the Friedrich Wilhelm Bessel Research Award from the Alexander von Humboldt Foundation of Germany, I had the luxury of working on the project full-time from January to April when I visited Munich University of Technology. I would like to thank Joachim Hagenauer and Ralf Koetter for nominating me for the award and for hosting my visit. I would also like to thank the Department of Information Engineering, The Chinese University of Hong Kong, for making this arrangement possible.

There are many individuals who have directly or indirectly contributed to this book. First, I am indebted to Toby Berger who taught me information theory and writing. I am most thankful to Zhen Zhang, Ning Cai, and Bob Li for their friendship and inspiration. Without the results obtained through our collaboration, the book cannot possibly be in its current form. I would also like to thank Venkat Anantharam, Vijay Bhargava, Dick Blahut, Agnes and Vincent Chan, Tom Cover, Imre Csiszár, Tony Ephremides, Bob Gallager, Bruce Hajek, Te Sun Han, Jim Massey, Prakash Narayan, Alon Orlitsky, Shlomo Shamai, Sergio Verdú, Victor Wei, Frans Willems, and Jack Wolf for their support and encouragement throughout the years. I would also like to thank all the collaborators of my work for their contribution and all the anonymous reviewers for their useful comments.

I would like to thank a number of individuals who helped in the project. I benefited tremendously from the discussions with David Tse who gave a lot of suggestions for writing the chapters on differential entropy and continuous-valued channels. Terence Chan, Ka Wo Cheung, Bruce Hajek, Siu-Wai Ho, Siu Ting Ho, Tat Ming Lok, Prakash Narayan, Will Ng, Sagar Shenvi, Xiang-Gen Xia, Shaohua Yang, Ken Zeger, and Zhixue Zhang gave many valuable comments at different stages of the writing. My graduate students Silas Fong, Min Tan, and Shenghao Yang proofread the chapters on network coding in great detail. Silas Fong also helped compose the figures throughout the book.

On the domestic side, I am most grateful to my wife Rebecca for her love. During our stay in Munich, she took good care of the whole family so that I was able to concentrate on my writing. We are most thankful to our family friend Ms. Pui Yee Wong for taking care of Rebecca when she was ill during the final stage of the project and to my sister Georgiana for her moral support. In this regard, we are indebted to Dr. Yu Lap Yip for his timely diagnosis. I would also like to thank my sister-in-law Ophelia Tsang who comes over during the weekends to help taking care of our daughter Shannon, who continues to be the sweetheart of the family and was most supportive during the time her mom was ill.