Shannon Theory 2.0

A New Shannon Theory Based on Algebraic Geometry

Fan CHENG
chengfan@sjtu.edu.cn
Information and Computing Lab
Department of Computer Science and Engineering
Shanghai Jiao Tong University
24/09/2022

Abstract

According to the result of 2022 Fields Medalist June Huh, as well as the Gaussian Completely Monotone Conjecture, differential entropy of heat flow admits an inner structure and can be further decomposed by a Hodge structure in Algebraic Geometry. It will fundamentally change Shannon Theory ever since 1948.

Agenda

- Background
- Gaussian Completely Monotone Conjecture
- The Application and Verification of GCMC
- GCMC and Hodge Theory
- Discussion and Summary

Background

Fundamental Tools

Before 1990

$$H(X, Y) \geq I(X; Y)$$
$$H(X) \geq H(X \mid Y)$$

Information Inequality

$$e^{2h(X+Y)} \geq e^{2h(X)} + e^{2h(Y)}$$

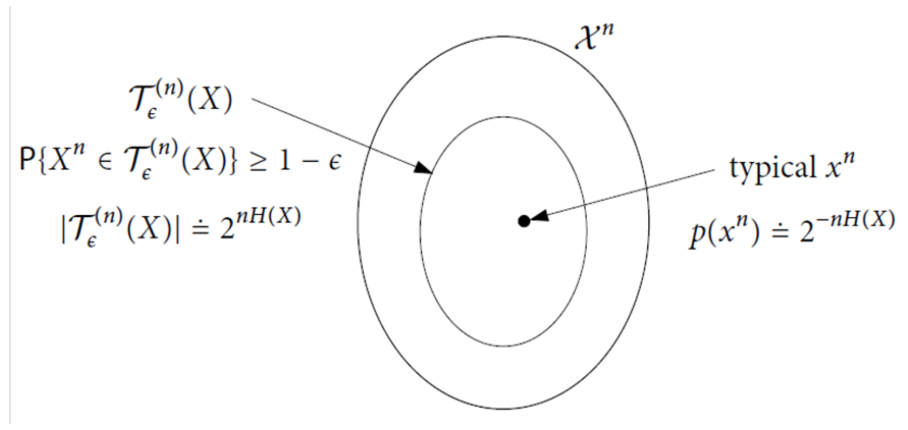
Entropy Power inequality

$$H(X \mid Y) \leq 1 + P_e \log |\mathscr{X}|$$

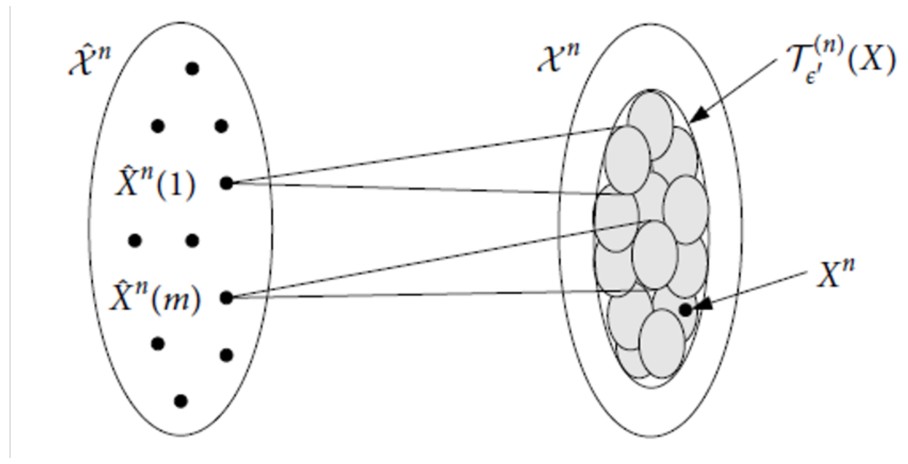
Fano's Inequality

$$\sum_{i=1}^{n} I(X_{i+1}^n; Y_i \mid Y^{i-1}, U) = \sum_{i=1}^{n} I(Y^{i-1}; X_i \mid X_{i+1}^n, U)$$

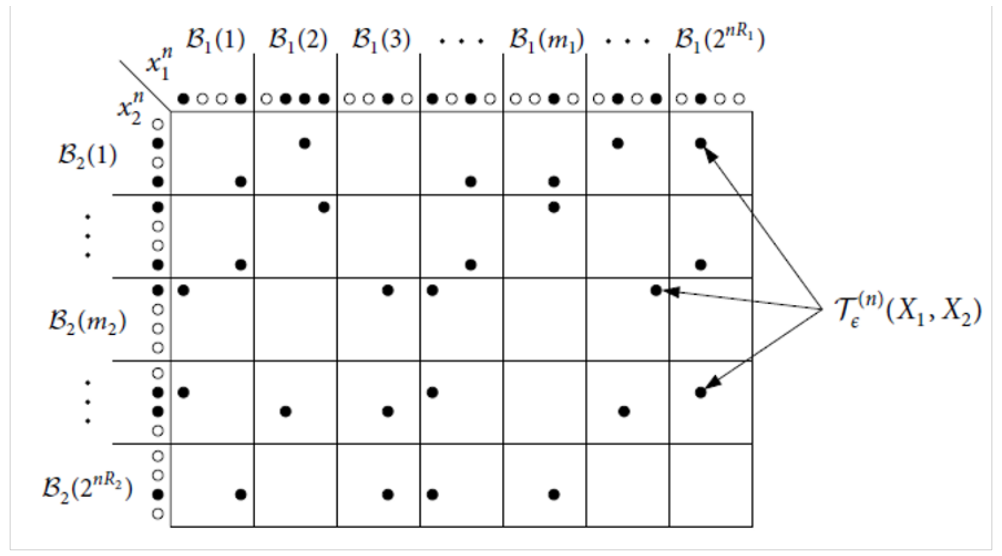
Csiszar Sum Identity



Typicality



Covering Lemma



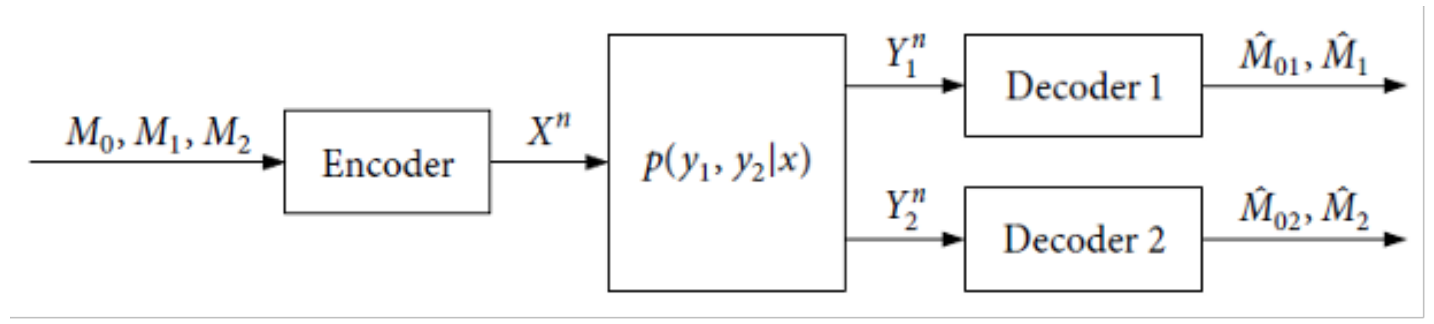
Random Binning

$$C_{SI-E} = C(P)$$

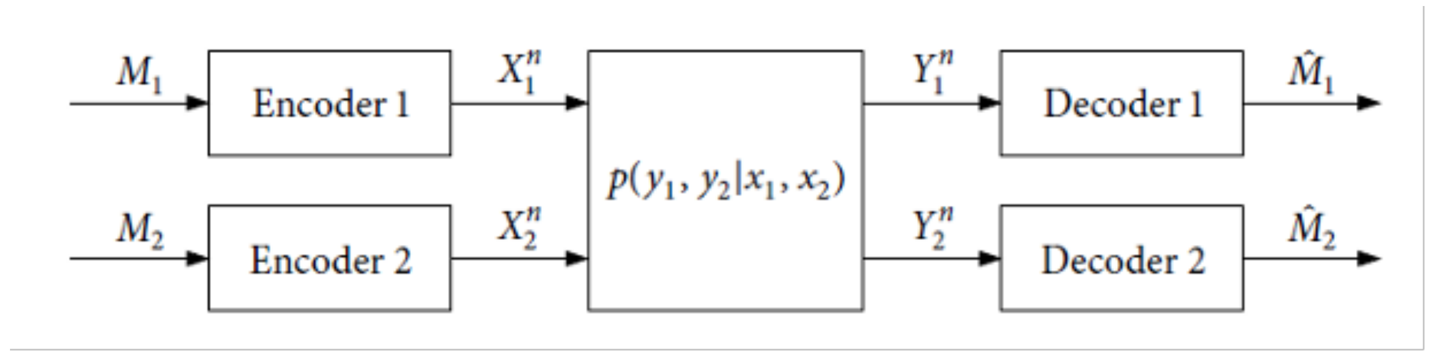
Dirty Paper Coding

Fundamental Problems

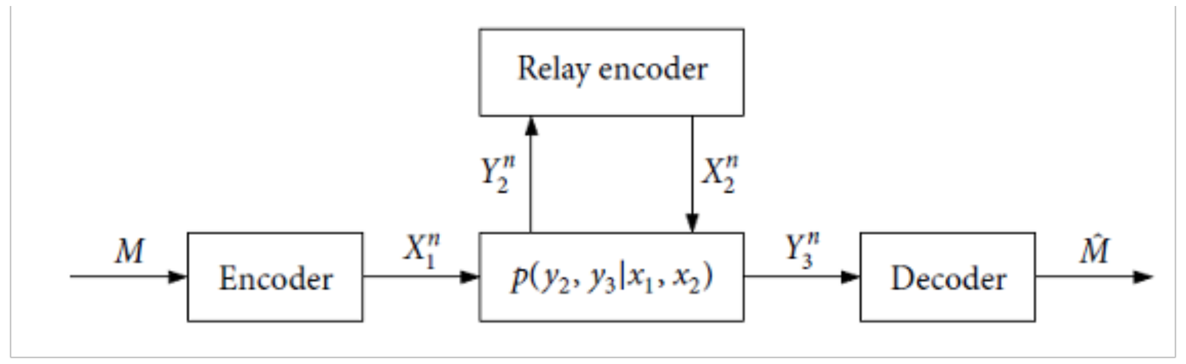
40-year Open Problems



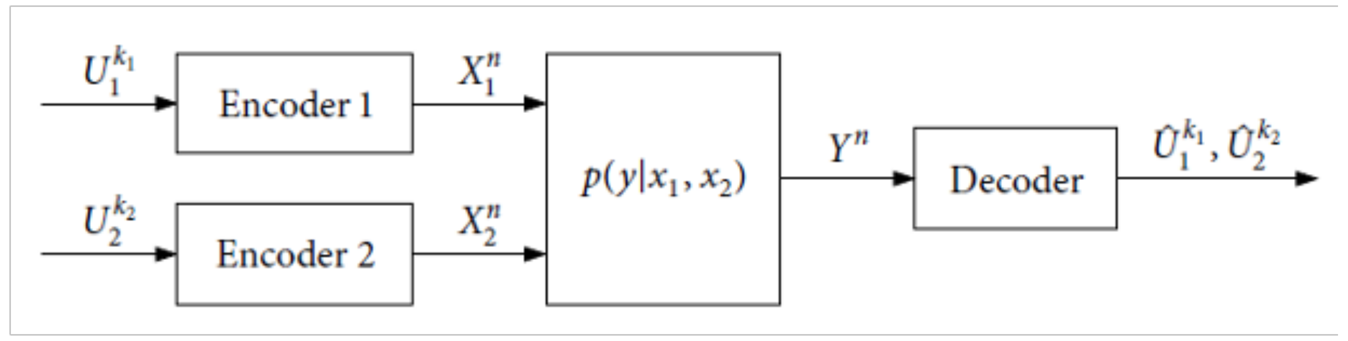
Broadcast Channel



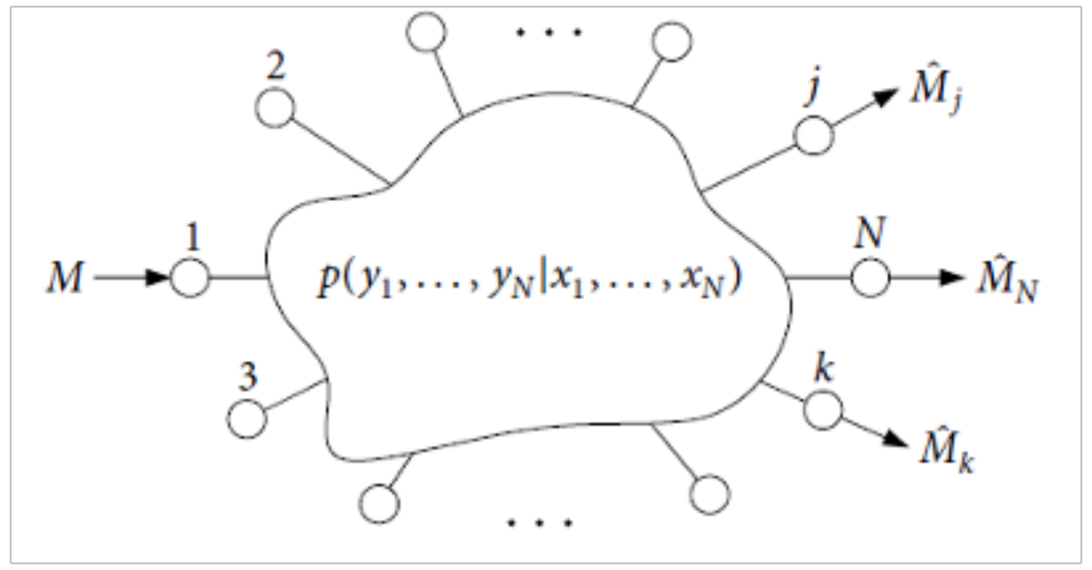
(Gaussian) Interference Channel



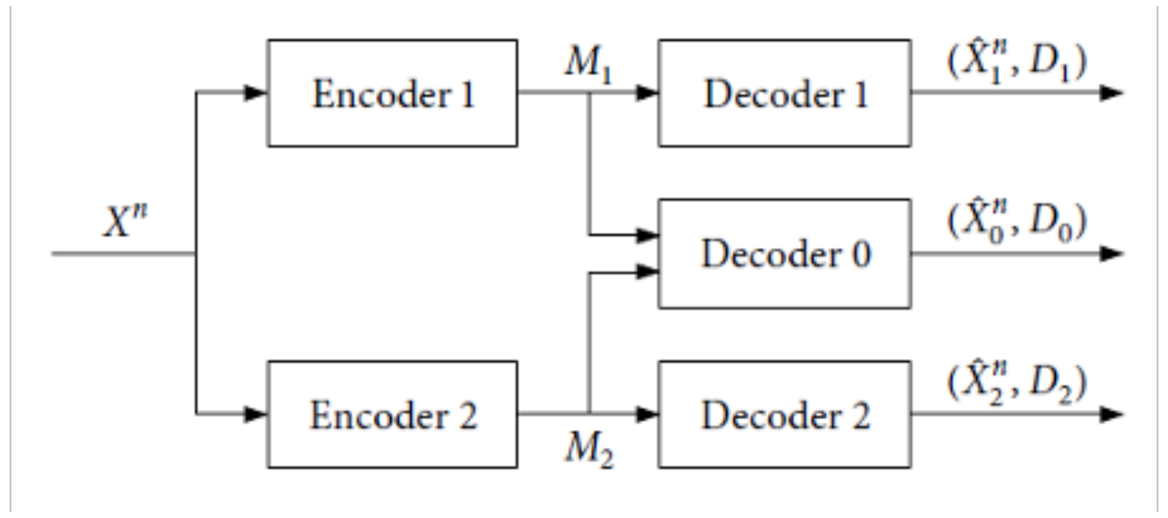
Relay Channel



Joint Multi-source channel coding



DMC-BC Network



Multiple Description Coding

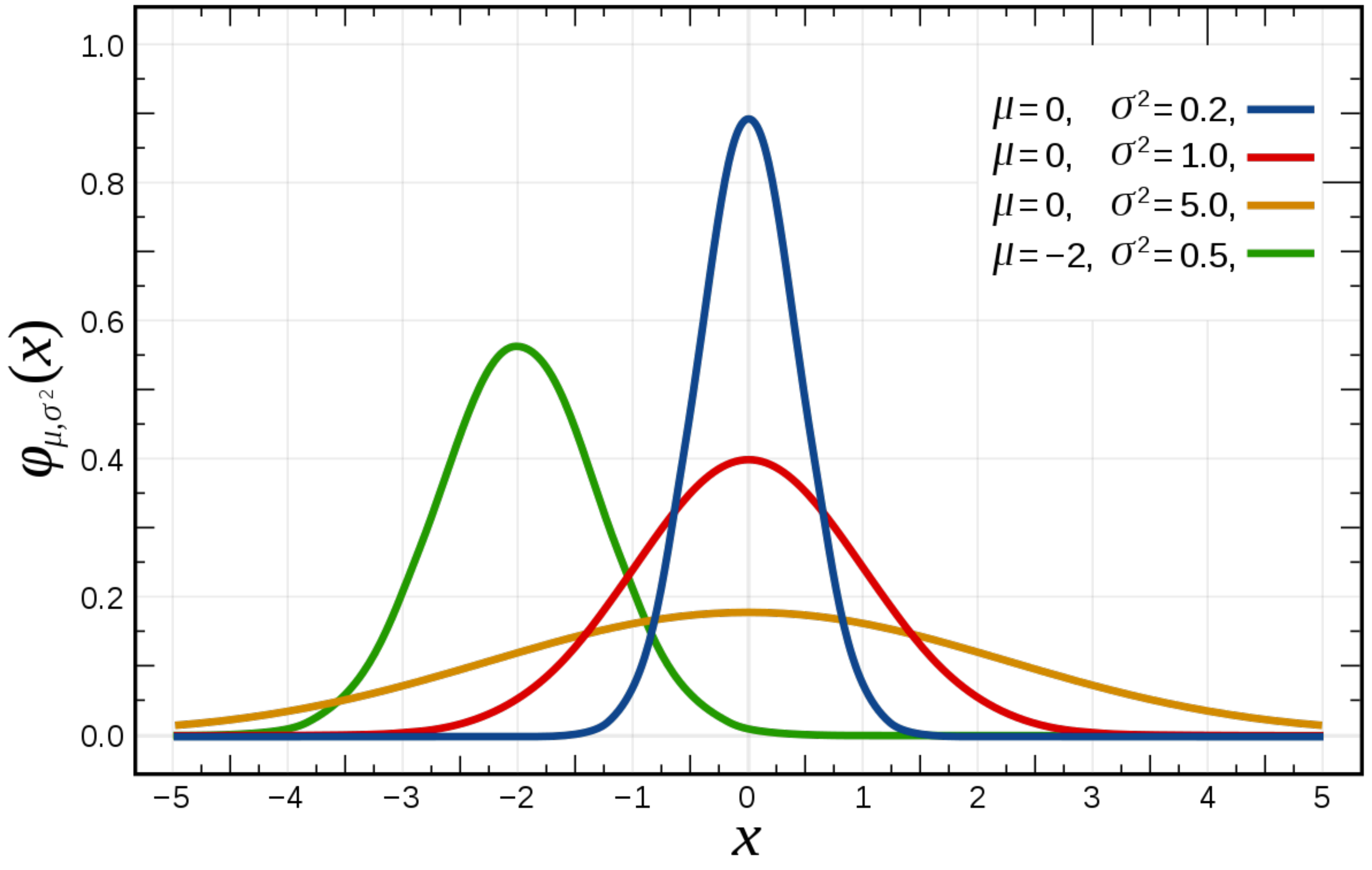
A. El Gamal and Y.-H. Kim, Network Information Theory, Cambridge Univ. Press, 2011.

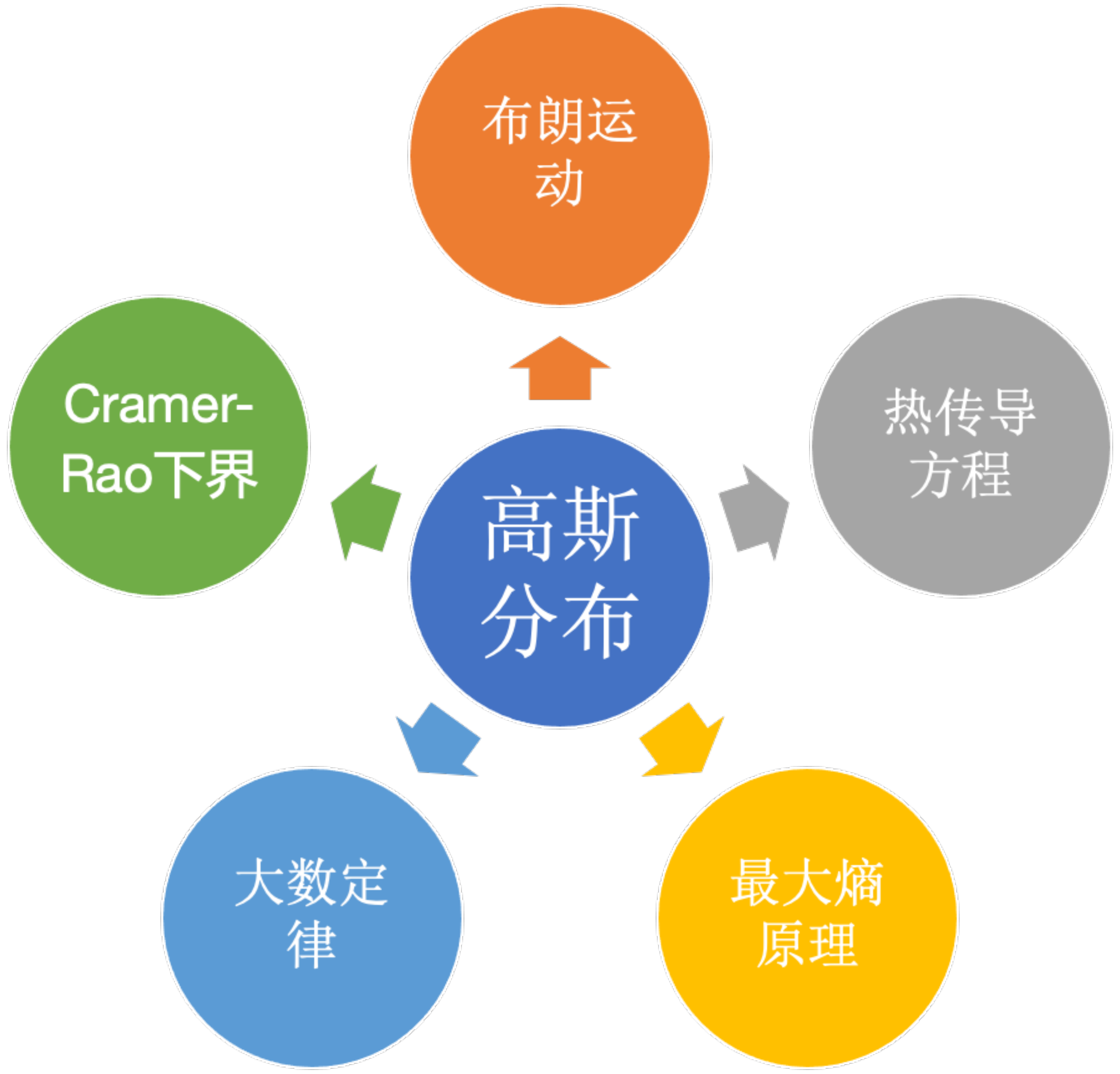
A New Mathematical Foundation

Gaussian Completely Monotone Conjecture

Gaussian Distribution

Fundamental Building-Block of Science and Engineering





Information theory of Gaussian Distribution

Entropy power inequality (EPI)

"Gaussian noise is the worst additive noise"

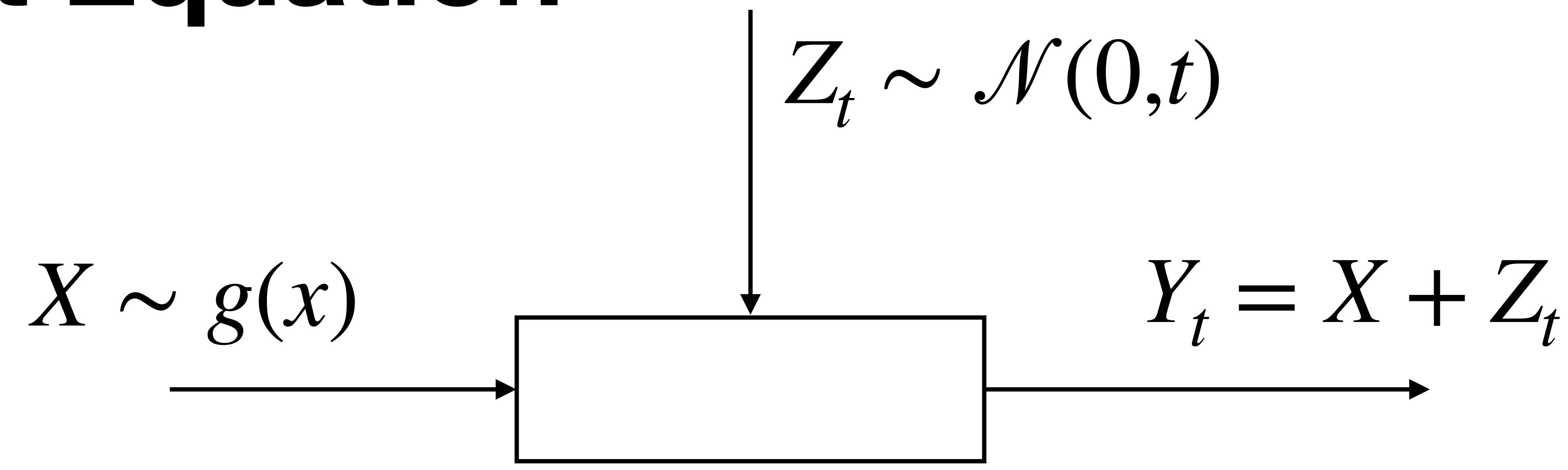
- (Shannon,1948) For any two independent continuous random variables X and Y

$$e^{2h(X+Y)} \geq e^{2h(X)} + e^{2h(Y)}$$

- The most important tool in Shannon Theory: uncertain principle, isoperimeter inequality
- Challenge： The channel capacity of Gaussian interference has not been settle down
- We have kept on studying EPI, but failed
  - Consensus： No new EPI and we need to find a way

Gaussian Channel

Equivalent to Heat Equation



- For any random variable X, which was affected by Z_t, the information received Y_t is the sum of X and Z_t
  - Y_t is referred to as the Gaussian mixed model (machine learning)
  - The p.d.f. of Y_t, $f(y,t)$ is the solution to heat equation

$$\frac{\partial}{\partial t} f(y,t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(y,t)$$

Differential entropy $h(Y_t) = -\int f(y,t) \log f(y,t) dy$ **is a basic quantity of information**

The derivatives of $h(Y_t)$

The meaning is not clear in information theory

$$h(Y_t) = \sum_i a_i t^i, \quad a_i \sim \frac{\partial^i}{\partial t^i} h(Y_t)$$

- $\frac{\partial}{\partial t} h(Y_t) = \frac{1}{2} I(Y_t) \geq 0$, I is the Fisher information
- $\frac{\partial^2}{\partial t^2} h(Y_t) \leq 0$
- $\frac{\partial^i}{\partial t^i} h(Y_t)$, $i \geq 3$ unknown

- My breakthrough (2013–2015)
  - $\frac{\partial^3}{\partial t^3} h(Y_t) \geq 0$
  - $\frac{\partial^4}{\partial t^4} h(Y_t) \leq 0$

Gaussian Completely Monotone Conjecture

The signed expression of the 3rd and 4th derivatives

Theorem 1: *For* $t > 0$,

$$\frac{\partial^3}{\partial t^3} h(Y_t) = \frac{1}{2} \int f \left(\frac{f_3}{f} - \frac{f_1 f_2}{f^2} + \frac{1}{3} \frac{f_1^3}{f^3} \right)^2 + \frac{f_1^6}{45 f^5} \mathrm{d}y.$$

Theorem 2: *For* $t > 0$,

$$\begin{aligned} &\frac{\partial^4}{\partial t^4} h(Y_t) \\ &= -\frac{1}{2} \int f \left(\frac{f_4}{f} - \frac{6}{5} \frac{f_1 f_3}{f^2} - \frac{7}{10} \frac{f_2^2}{f^2} + \frac{8}{5} \frac{f_1^2 f_2}{f^3} - \frac{1}{2} \frac{f_1^4}{f^4} \right)^2 \\ &\quad + f \left(\frac{2}{5} \frac{f_1 f_3}{f^2} - \frac{1}{3} \frac{f_1^2 f_2}{f^3} + \frac{9}{100} \frac{f_1^4}{f^4} \right)^2 \\ &\quad + f \left(-\frac{4}{100} \frac{f_1^2 f_2}{f^3} + \frac{4}{100} \frac{f_1^4}{f^4} \right)^2 \\ &\quad + \frac{1}{300} \frac{f_2^4}{f^3} + \frac{56}{90000} \frac{f_1^4 f_2^2}{f^5} + \frac{13}{70000} \frac{f_1^8}{f^7} \mathrm{d}y. \end{aligned}$$

Gaussian Completely Monotone Conjecture

Related publications 2013—2022

- F. Cheng, "Generalization of Mrs. Gerber's Lemma," Communications in Information and Systems, vol. 14, no. 2, pp. 79-86, 2014 （work finished in 2011-2012，published in 2014）
- F. Cheng, "Some conjecture on Entropy Power inequality," 2013 Workshop on Coding and Information Theory, HKU, Dec. 2013
- F. Cheng and Y. Geng, "Convexity of Fisher Information with Respect to Gaussian Perturbation," 2014 Iran Workshop on Communication and Information Theory, (IWCIT 2014)
- F. Cheng and Y. Geng, "Higher Order Derivatives in Costa's Entropy Power Inequality," IEEE Transactions on Information Theory, vol. 61, no. 11, pp. 5892-5905, Nov. 2015
- F. Cheng, "How to Solve Gaussian Interference Channel," The 2019 Workshop on Probability and Information Theory (WPI 2019), HKU, Aug. 2019
- F. Cheng, "A Reformulation of Gaussian Completely Monotone Conjecture: A Hodge Structure on the Fisher Information along Heat Flow," https://arxiv.org/abs/2208.13108

Gaussian Completely Monotone Conjecture

2013—2015

Summary: The signs of derivatives of $\frac{\partial^i}{\partial t^i} h(Y_t)$ is +, -, +, -, ?, ?, ?

Conjecture: Signs alternate in + and -

- (Conjecture 1) For any random variable X, the sign of $\frac{\partial^i}{\partial t^i} h(Y_t)$
  - When i is even, it is negative
  - When i is odd, it is positive
- (Conjecture 2) $I(Y_t)$ is log-convex in t

Breakthrough since 1966

Problems in C. Villani's Textbook：McKean 1966, CM functions

C. Villani
2010 Fields

A review of mathematical topics in collisional kinetic theory

Cédric Villani

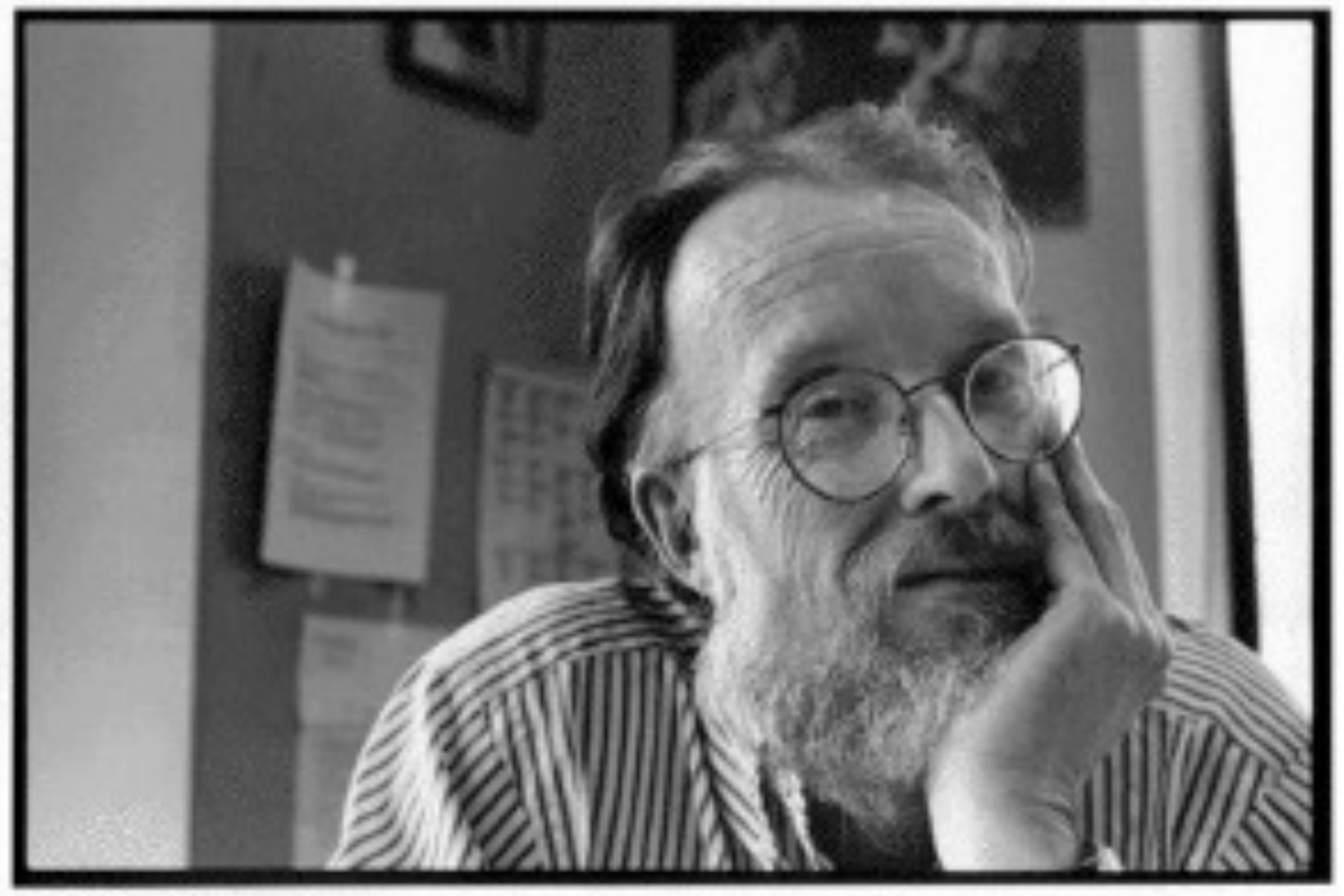
completed: October 4, 2001
revised for publication: May 9, 2002
most recent corrections: June 7, 2006

Cedric Villani <villani@ihp.fr>
to me

Dear Fan,
Thank you for your message. Sorry for not replying earlier.
I am not aware of a result like this. I suggest that you check with Amir Dembo or the Cover-Thomas duo. Good luck!
Best, Cedric

There is a history of conjectures of "complete monotonicity" of functionals in the context of the Boltzmann equation, which has turned to be quite wrong. You can see some references in my online review on the Boltzmann equation, see the chapter about "Maxwellian collisions" and the "McKean conjecture", which turned out to be false. See also p.166 of my review (I attach a version). In view of that story, I think Conjecture 1 is too daring; just having n=1 and n=2 is not convincing enough to give a hint of it. (Please check whether it is not explicitly the 2nd McKean conjecture; it has been a long time and I don't remember details well.)

Best, Cedric



H. P. McKean
NYU

Completely Monotone

Hausdorff—Bernstein—Widder 1920s

F. Hausdorff

S. N. Bernstein

D. Widder

- Example： The signs of the derivatives of $1/t$ are -, +, -, +, ……
- A function $f(t)$ is called completely monotone (CM) if its derivatives alternate in signs
  - It is trivial to conduct derivatives but it is not easy to show the signs

Gaussian Completely Monotone Conjecture (GCMC)

1. The Fisher information $I(Y_t)$ is CM
2. $I(Y_t)$ is log-convex

Two facts on CM functions

H.B.W. 1920

1. $f(t)$ is CM, then $f(t)$ is log-convex

- Gaussian Completely Monotone Conjecture

  1.Fisher information $I(Y_t)$ is CM

  2.$I(Y_t)$ is log-convex

- Conjecture 1 implied Conjecture 2

2.Laplace Representation of CM functions

- $f(t)$ is CM, iff there exists a non-decreasing Borel measure $\mu(x)$ on $[0, +\infty)$ such that

$$f(t) = \int_x e^{-xt} d\mu(x)$$

New Expression for Information

GCMC looks great, but what's its applications?

The Application and Verification of GCMC

The mathematical meaning of CM functions

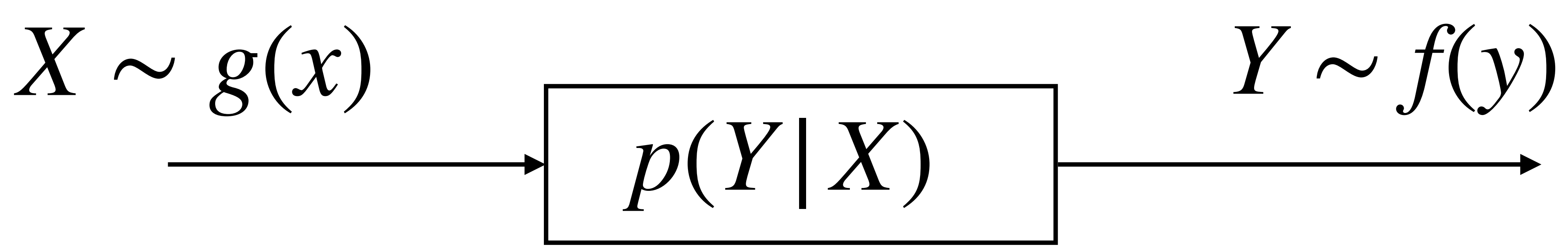
HKU2019： Information is decomposable and reversable

$$I(Y_t) = \int_x e^{-xt} d\mu(x)$$

- The decomposition of Fisher information
  - e^{-xt} is for CM
  - $\mu(x)$ is the identity of X, regardless of t
- Laplace transformation is reversible
  - By $Y_t = X + Z_t$, though Z_t has disturbed X, $\mu(x)$ remained unchanged
  - Recall that, in the second law of thermodynamics, the status of the system is not reversible

The information meaning of CM functions

It is up to its mathematical meaning



- X is transformed into Y by the channel
- The relation between X and Y is determined by $p(Y|X)$
- So far, no theory to govern X and Y
  - The problem is intractable if we add even one more node in the point to point case
- The technical reason why network information theory is always hard

$$I(Y_t) = \int_x e^{-xt} d\mu(x)$$

gives a constraint on the Gaussian node：

reversible and decomposable

Potential Application: Gaussian interference channel

An Application in Gaussian Multiuser Channel

2019-2021

Log-convexity of Fisher information along heat flow

Michel Ledoux
Institut de Mathématiques de Toulouse
University of Toulouse – Paul-Sabatier
Toulouse, France
Email: ledoux@math.univ-toulouse.fr

Chandra Nair and Yan Nan Wang
Dept. of Information Engg.
The Chinese University of Hong Kong
Sha Tin, N.T., Hong Kong
Email: {chandra,dustin}@ie.cuhk.edu.hk

Michel Ledoux

Chandra Nair

Yan Nan Wang

- This paper establishes the log-convexity of Fisher information for scalar random variables along the heat flow, thus resolving a conjecture posed in [1]
- Such results may also be useful in showing the uniqueness of local maximizers in such settings as is observed in settings such as the MIMO Gaussian broadcast channels

Verification

$I(Y_t)$ is log-convex

- Ledoux proved that $I(Y_t)$ is log-convex
- Gaussian Completely Monotone Conjecture (GCMC)

  1. The Fisher information $I(Y_t)$ is CM

  2. $I(Y_t)$ is log-convex

  - C1 implies C2

Imagine: C2 is merely a point of C1!!!

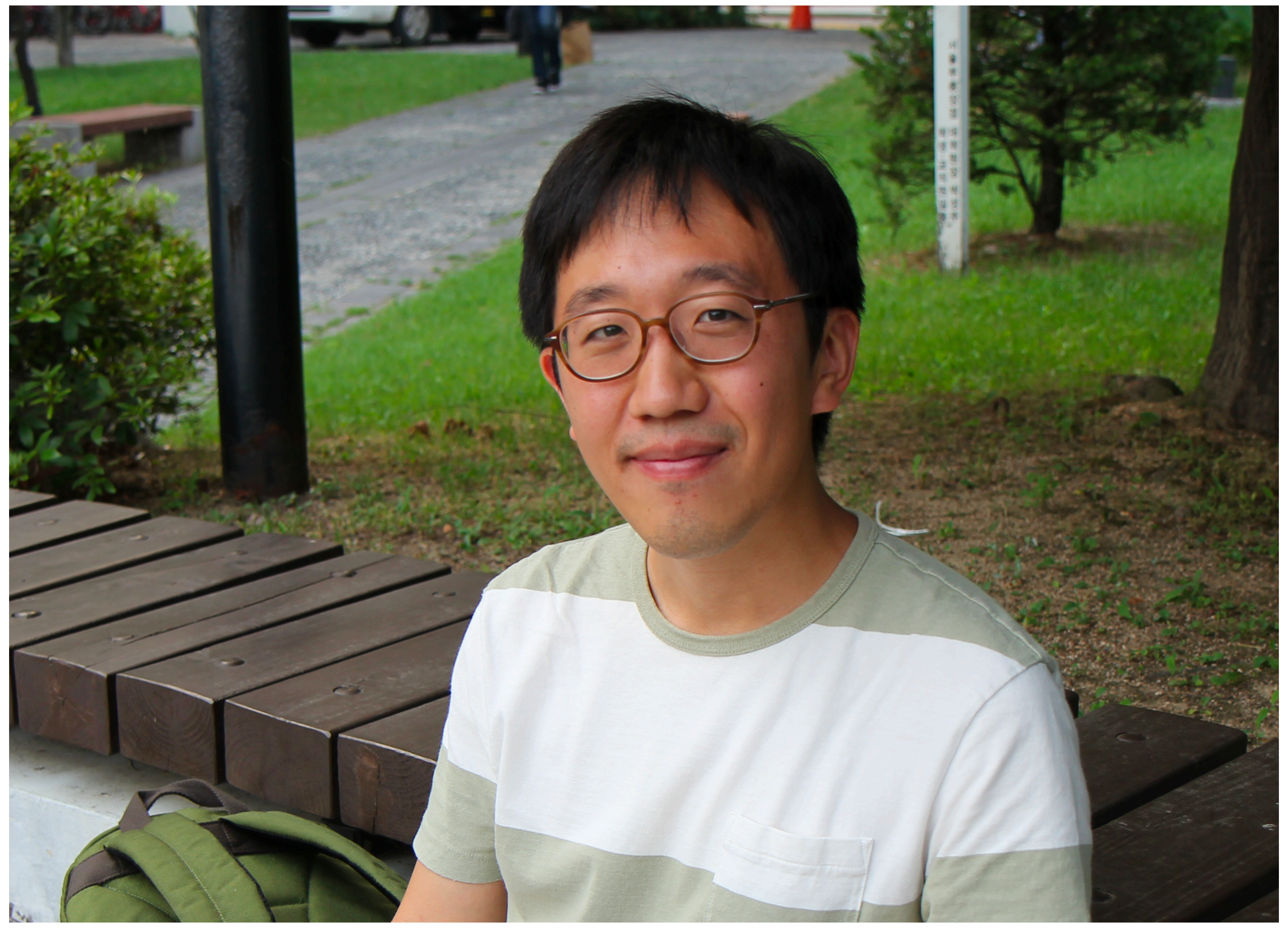
C2: A big surprise later

- The proof of C2 provides a necessary condition of C1

GCMC and Hodge Theory

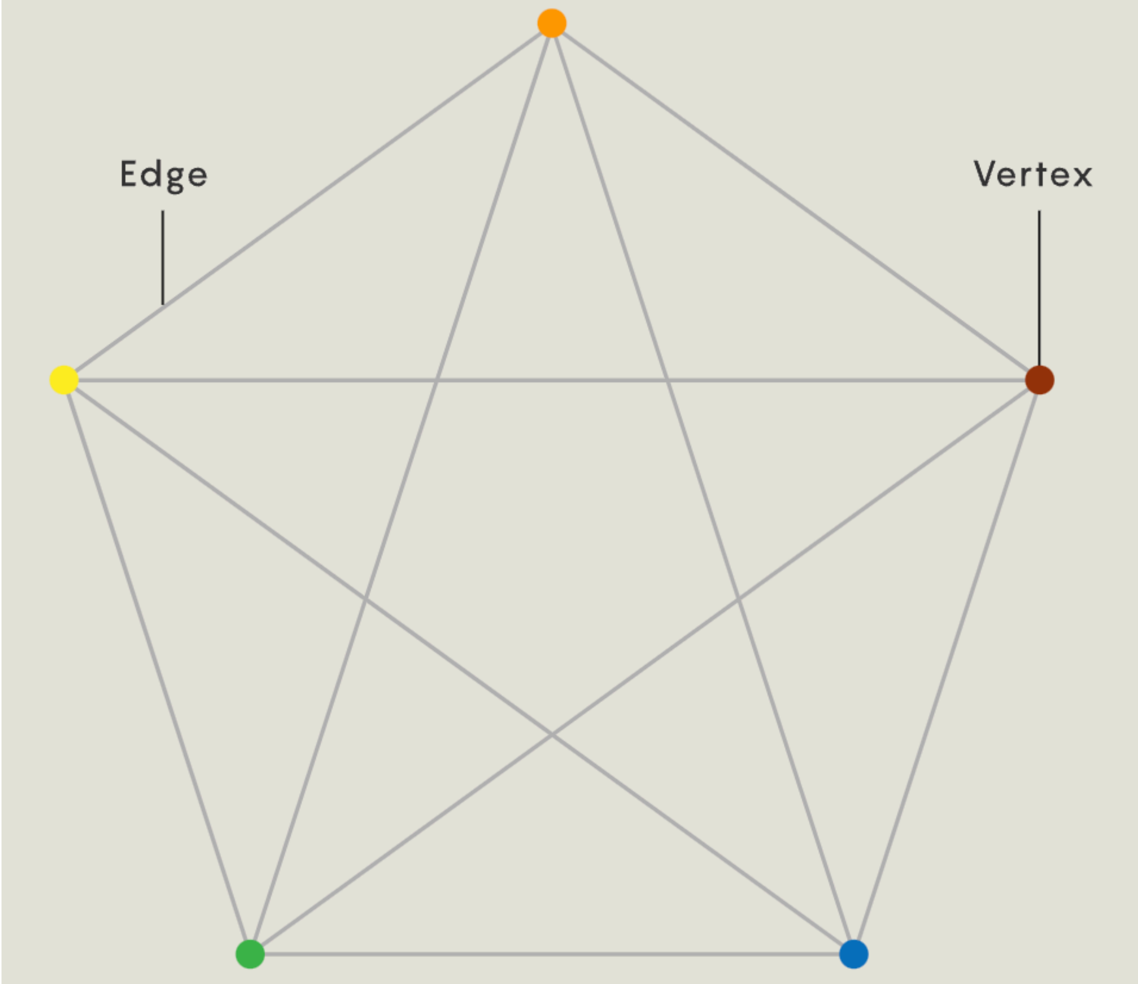
2022 International Congress of Mathematicians

2022.07.06—2022.07.14



June Huh（许埈珥）
2022 Fields Medal

- Four Color Theorem（machine proof)
- Chromatic polynomial
- C: the number of colorings with n colors
  $C = n^5 - 10n^4 + 35n^3 - 50n^2 + 24n$
- The coefficients $\{a_i\}$ of C above satisfies that $a_i^2 \geq a_{i-1}a_{i+1}$
- If $\{a_i\}$ satisfies that $a_i^2 \geq a_{i-1}a_{i+1}$, then it is a log-concave sequence
- **Conjecture: The coefficients of C form a log-concave sequence for any C**



https://www.quantamagazine.org/june-huh-high-school-dropout-wins-the-fields-medal-20220705/

Constructive proof of log-concave sequence

Solved several long-standing open problems of log-concave sequence

- June Huh used Hodge theory to study combinatorics
  - Construct Complex algebraic variety, study its homology and cohomology
- Open problem 1： chromatic polynomial
- Open problem 2： matroid
- Open problem 3： geometry lattice

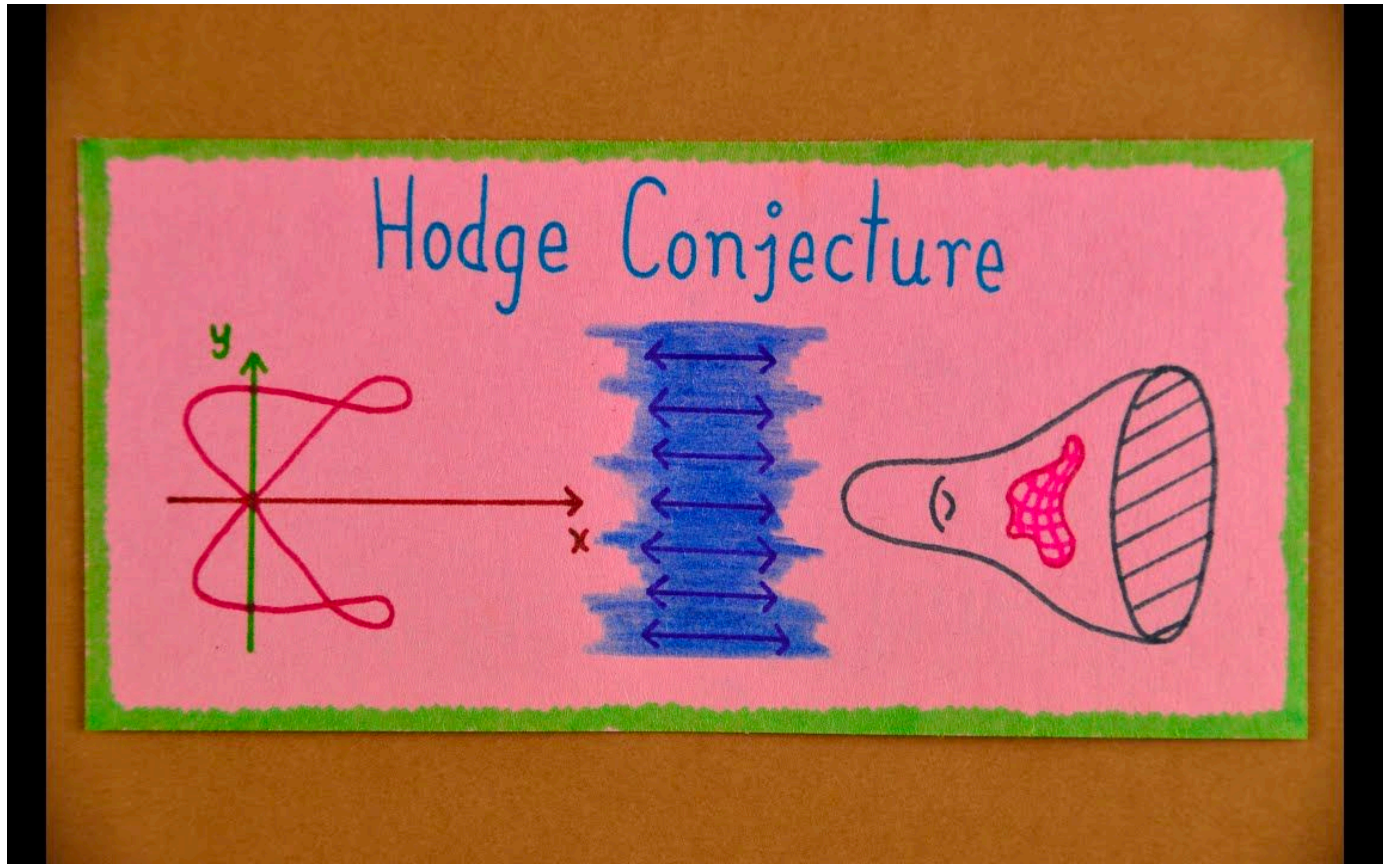
"bringing the ideas of Hodge theory to combinatorics, the proof of the Dowling–Wilson conjecture for geometric lattices, the proof of the Heron–Rota–Welsh conjecture for matroids, the development of the theory of Lorentzian polynomials, and the proof of the strong Mason conjecture"—Fields Medal Citations

Hodge Theory

Central to contemporary math

- Hodge conjecture: It is one of the seven Millennium Prize Problems set up by the Clay Mathematics Institute.

Millennium Prize Problems
• Birch and Swinnerton-Dyer conjecture • **Hodge conjecture** • Navier–Stokes existence and smoothness • P versus NP problem • Poincaré conjecture (solved) • Riemann hypothesis • Yang–Mills existence and mass gap



CM Functions and Hodge Theory

Linked with Algebraic Geometry

- Hausdorff et. al. showed that, if $f(t)$ is CM, then $f(t)$ is log-convex
- $(\log f(t))'' \geq 0 \rightarrow \dfrac{f''f - (f')^2}{f^2} \geq 0 \rightarrow f''f \geq (f')^2$

  f, f', f'' is a log-convex sequence
- A new characterization of CM functions: if $f(t)$ is CM, then

  $$f, f^{(1)}, f^{(2)}, \ldots, f^{(n)}, \ldots$$

  is a log-convex sequence.
- If $\{a_i\}$ is a log-convex sequence, then $\{1/a_i\}$ is a log-concave sequence

Idea: June Huh's method→CM functions

- My intuition: why it works

June Huh’s summary

It is a general method for log-concave sequence

> I believe that behind any log-concave sequence that appears in nature, there is such a “Hodge structure” responsible for the log-concavity.
>
> June Huh

- “Tropical geometry of matroid,” June Huh
- “Hodge Theory of Matroids,” Karim Adiprasito, June Huh, and Eric Katz

The derivatives of $h(Y_t)$

Meaning of information theory

$$h(Y_t) = \sum_i a_i t^i, \quad a_i \sim \frac{\partial^i}{\partial t^i} h(Y_t)$$

- Gaussian Completely Monotone Conjecture (GCMC)

  1. The Fisher information $I(Y_t)$ is CM

  2. $I(Y_t)$ is log-convex

Entropy（bit）can be further decomposed and it has an inner structure （e.g., atoms and quarks）

- $\{\frac{\partial^i}{\partial t^i} h(Y_t), i = 1,2,\ldots\}$ is supported by a Hodge structure

- Hodge theory is merely a branch of the grand algebraic geometry family
- GCMC is one of the hard problems in information theory

Reshape information theory via algebraic geometry

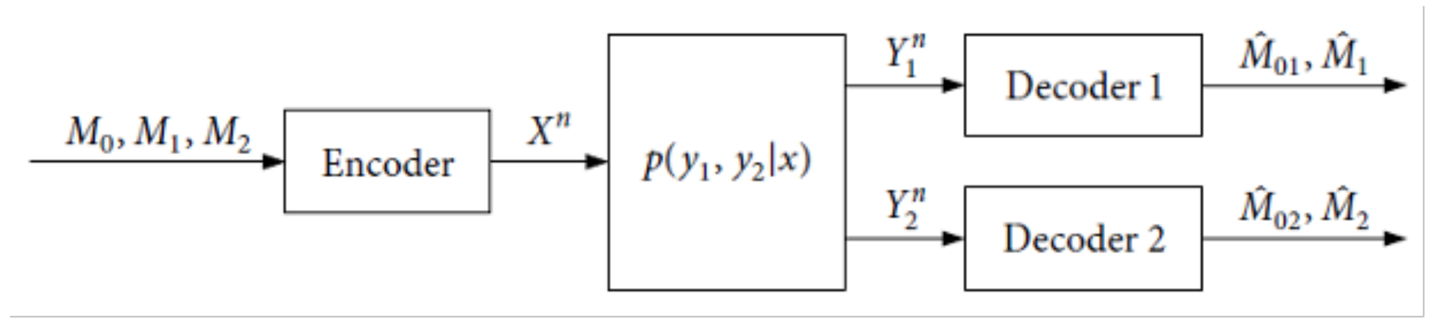
Discussion and Summary

- A new mathematical foundation of Shannon Theory
- Some fundamental open problems may be solved

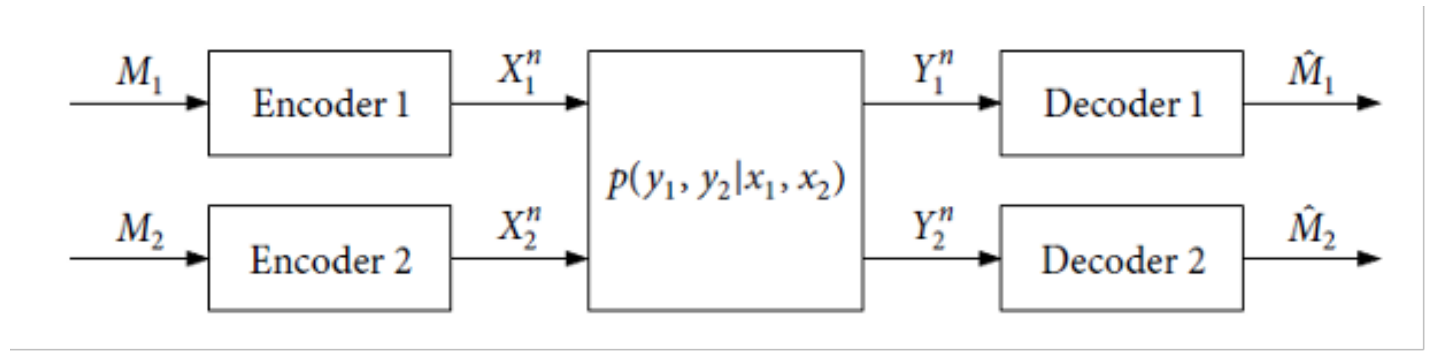
How hard is information theory

Very close to some hard problems in AG

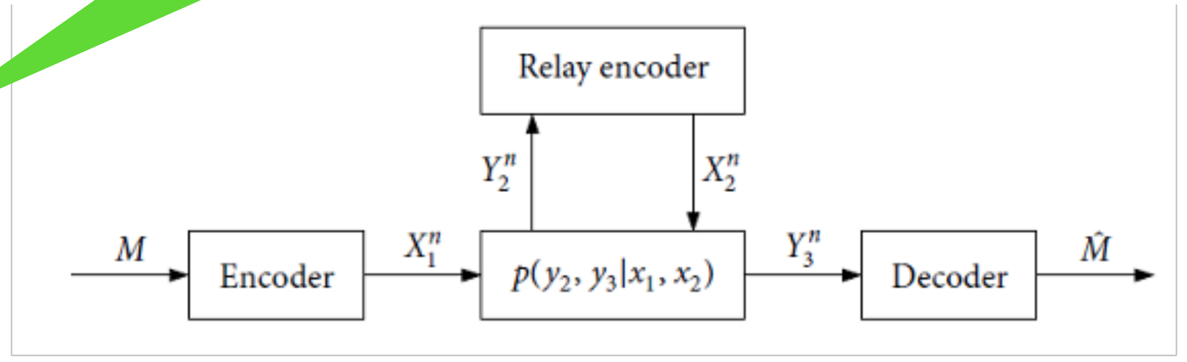
We have done our best :)



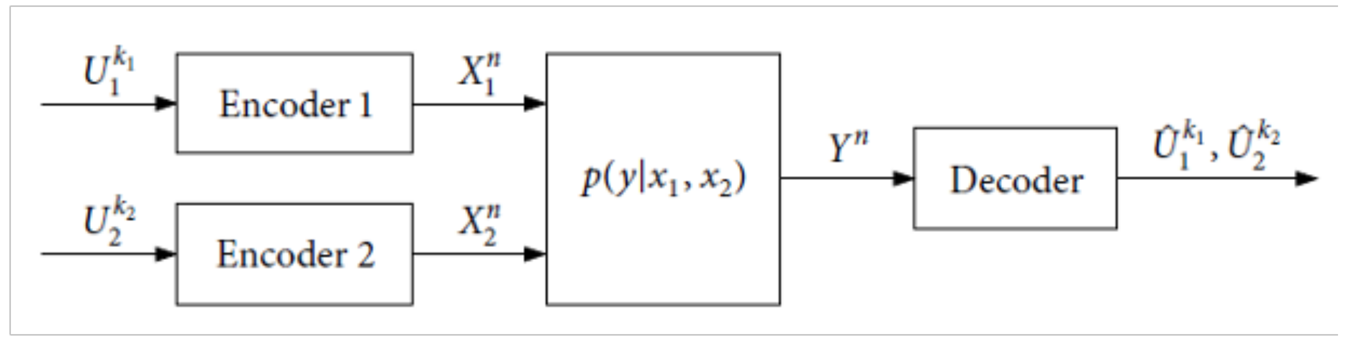
Broadcast Channel



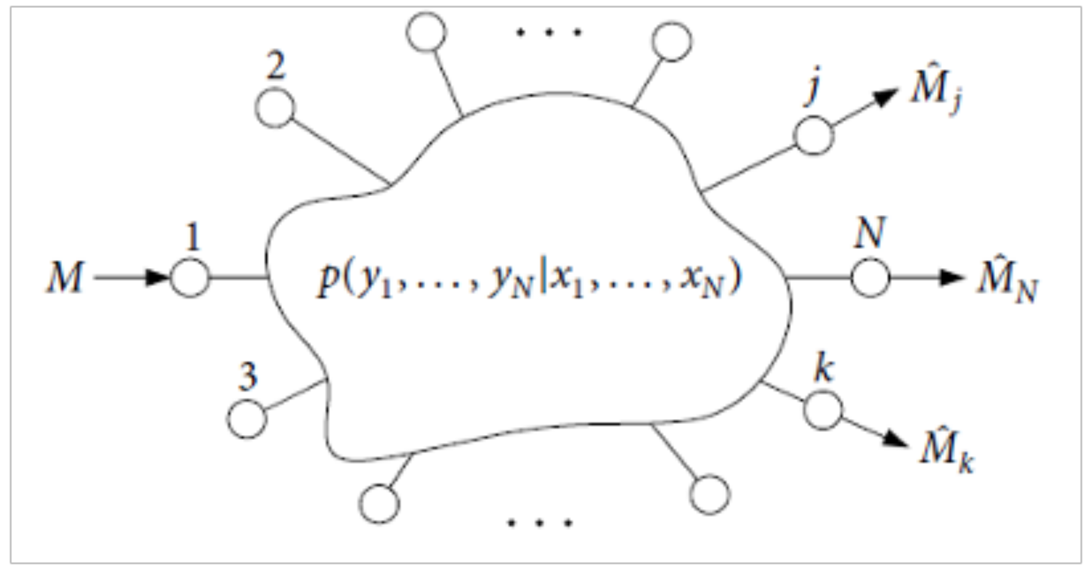
(Gaussian) Interference Channel



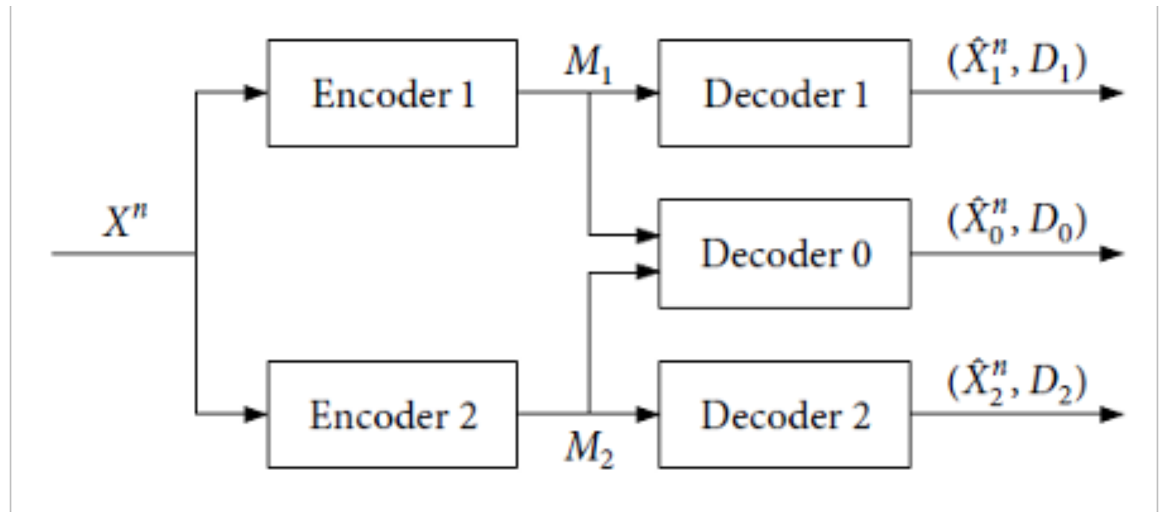
Relay Channel



Joint Multi-source channel coding



DMC-BC Network



Multiple Description Coding

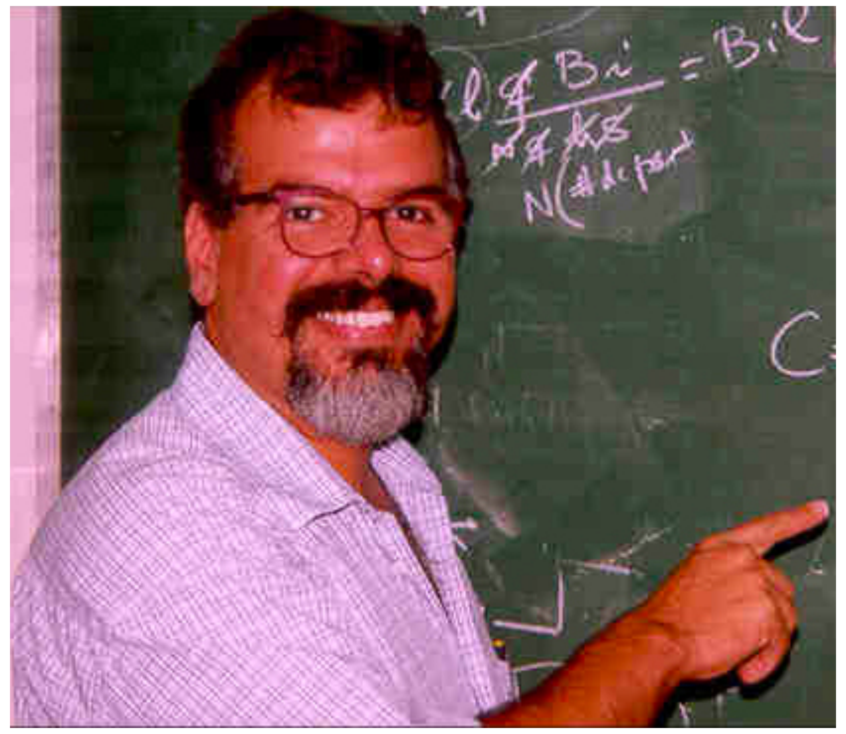
Reshape information theory via algebraic geometry (IT2.0) has already on its way

Will be some fundamental change in 3 years


C. E. Shannon


T. Cover

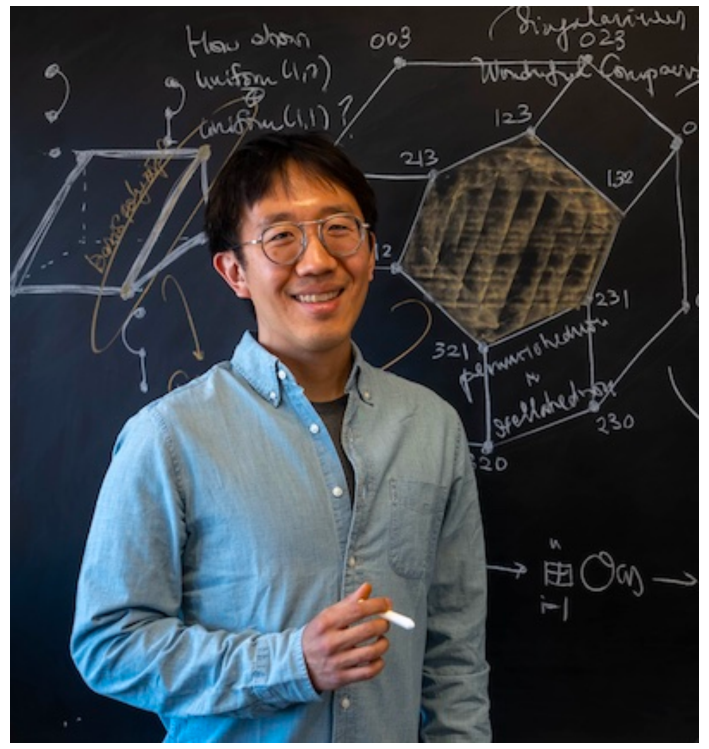

A. El Gamal


M. H. Costa


R. W. Yeung


F. Hausdorff


C. Villani


J. Huh

All the past is a prelude
The spirit of fundamental research

Thanks!

A Reformulation of Gaussian Completely Monotone Conjecture:
A Hodge Structure on the Fisher Information along Heat Flow
https://arxiv.org/abs/2208.13108
https://ichengfan.github.io/IT/