信息论2.0

基于代数几何的新一代香农信息理论

程帆， 博士、副教授
chengfan@sjtu.edu.cn
信息与计算实验室，计算机科学与工程系，上海交通大学
2022年09月26日

简　介

根据2022年Fields奖得主June Huh的工作，以及高斯分布完全单调猜想，信息熵具有特殊的内部结构，可以进一步分解，其分解由代数几何Hodge结构所约束。这将从根本上改变自1948年以来香农信息论的研究

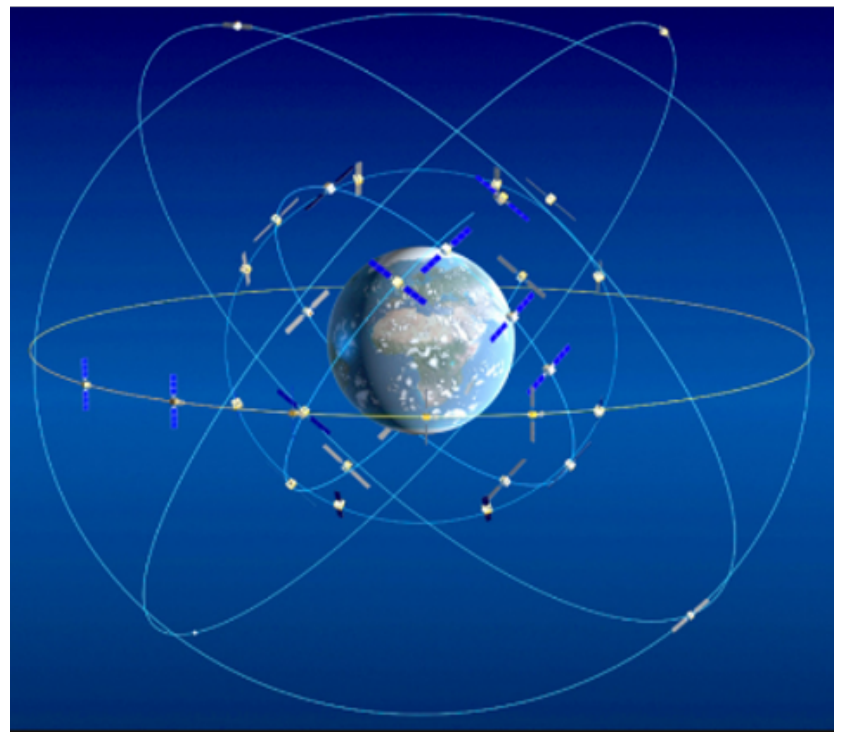
报告安排

- 背景介绍
- 高斯分布的完全单调猜想
- 完全单调猜想的应用以及在干扰信道的验证
- 完全单调猜想与Hodge结构
- 讨论与总结

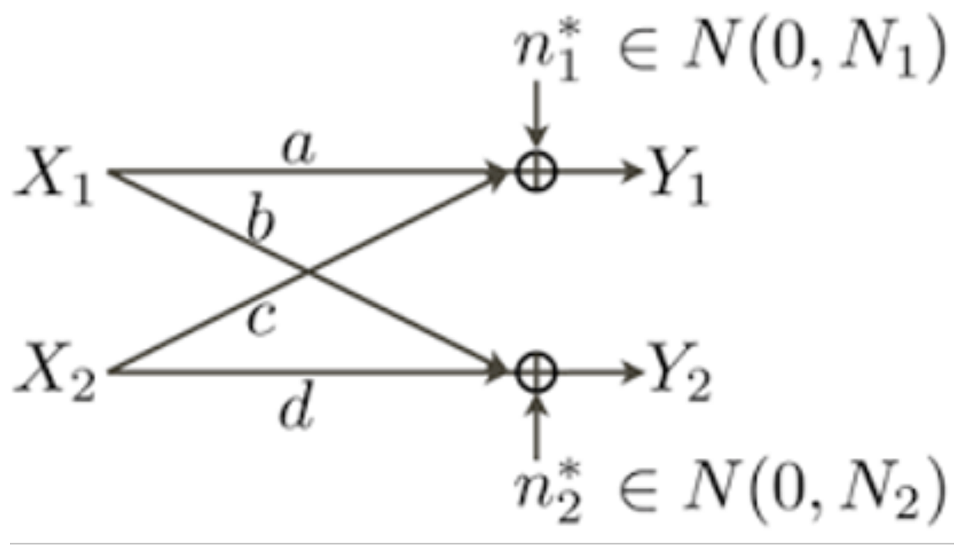
背景介绍

香农理论（信息理论）

香农理论研究why，编码理论研究how



实际通信系统



合理的数学抽象

$$C \le \max I(X;Y)$$

$$C = \frac{1}{2}\log(1 + \frac{P}{Q})$$

严格的数学定理

- 香农理论: C. E. Shannon于1948年创建，主要研究通信系统的基本极限
  - 信源压缩与信道传输
  - 信息科学的数学基础
- 误解：
  - 香农理论是严格的数学理论，不可能被颠覆
  - 香农极限是数学界限，不可能被突破
- 可以重新定义问题，改变模型假设，获得新的极限

信息论特指香农理论

编码理论另说

信息理论提供指引

固态电子技术提供动力

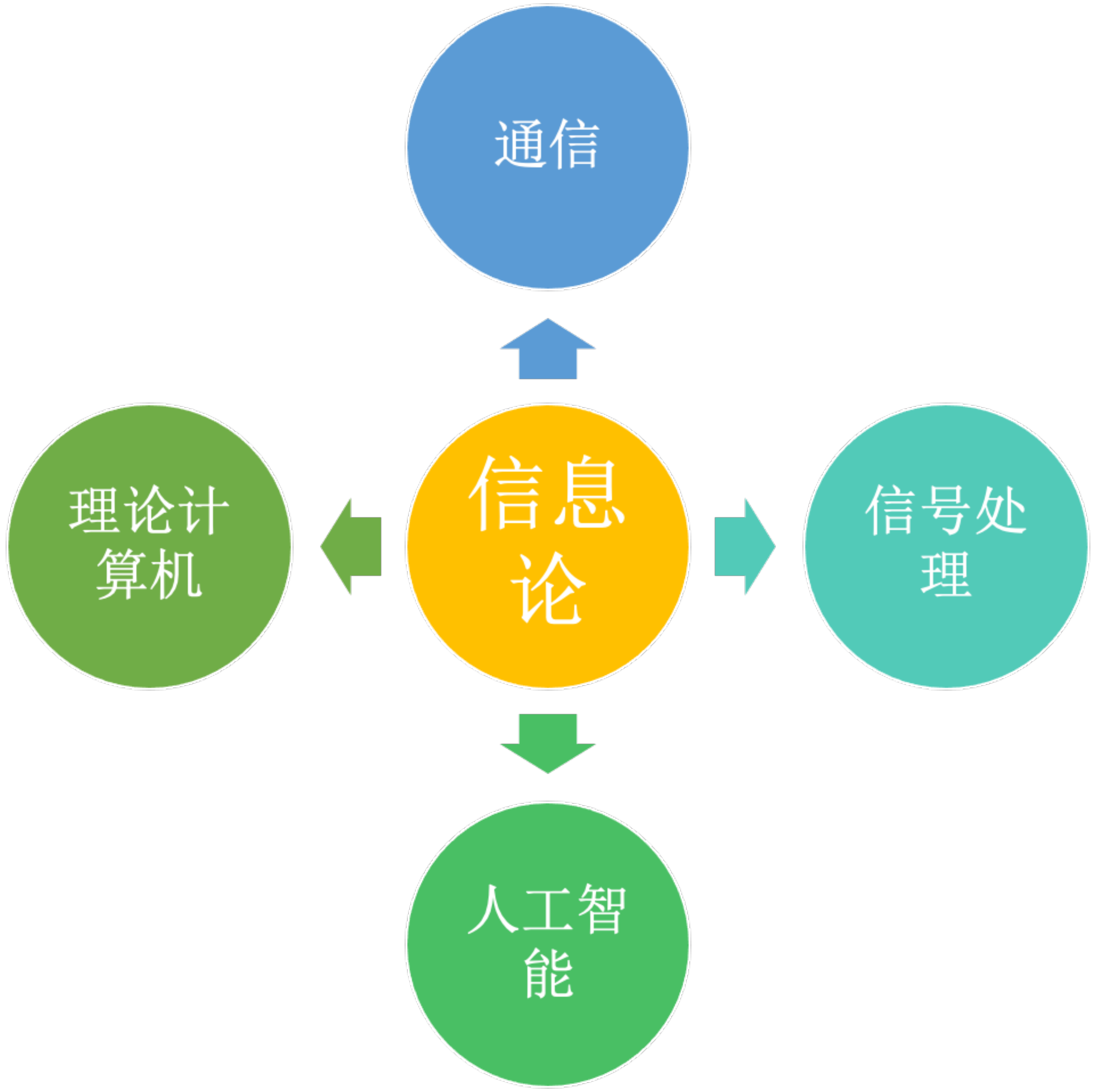
“Advances in digital communication in the latter half of this century were guided by the lessons of information theory but fueled by the progress in solid state electronics.”

IEEE Communications Magazine, Sept. 1991, A. J. Viterbi

低维情况下，可以通过经验或者常识来构造。但是对于复杂的高维情况，人的常识和经验往往是错误的，需要科学理论来指导

信息论对各领域的支撑

各领域的数学基础：信息的数学定义和基本规律



香农理论面临的挑战

打造新的数学基础工具

大规模、异步、高码率、低延时的通信系统的信息理论?

- 香农理论的数学基础，在1948年Shannon奠定以后，变化不大
  - 信息论的数学基础（数学基础的数学基础）很难改进与创新
  - 太稳定难以应付新的问题：复杂信息模型的挑战
- 改进和打造新的信息论数学基础工具
  - 没有太多新的进展，why?

基本数学工具

大部分是1990年前提出

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

信息不等式

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

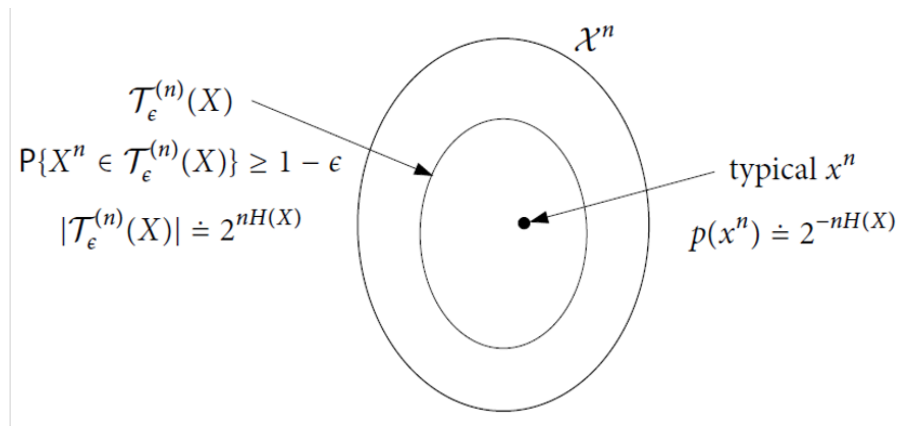
熵幂不等式

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

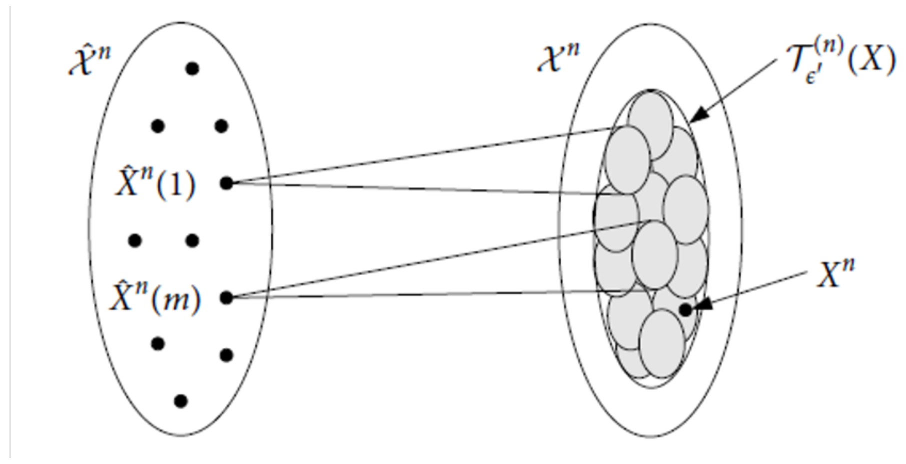
Fano不等式

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

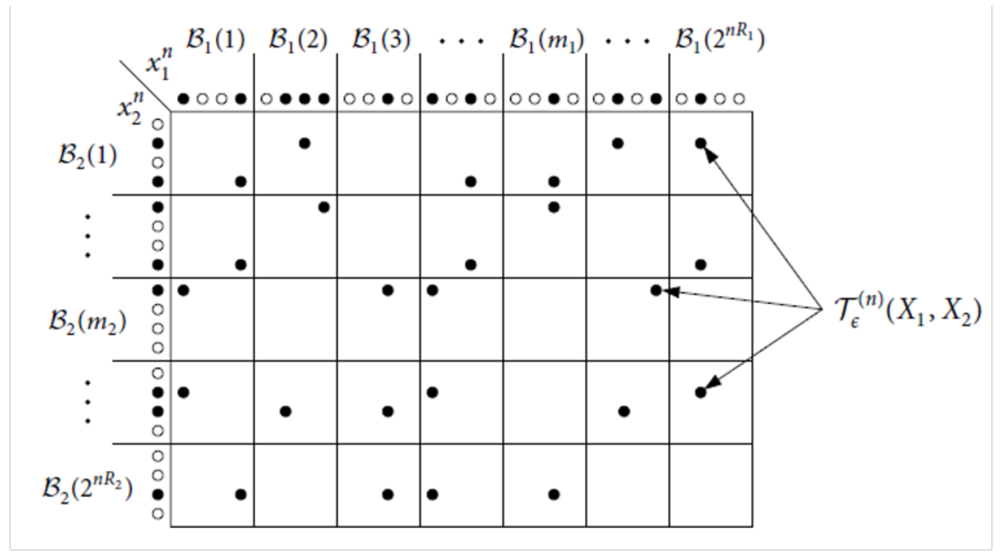
Csiszar和恒等式



典型性



覆盖引理



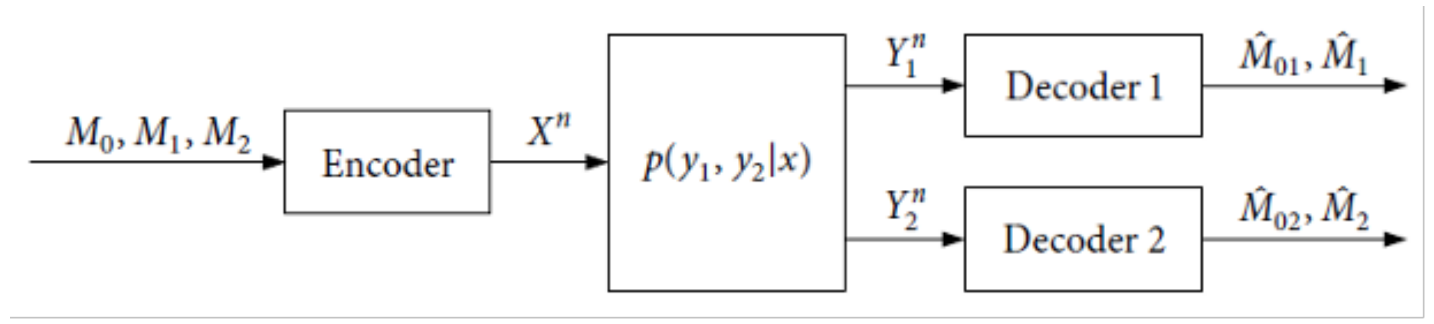
随机装箱编码

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

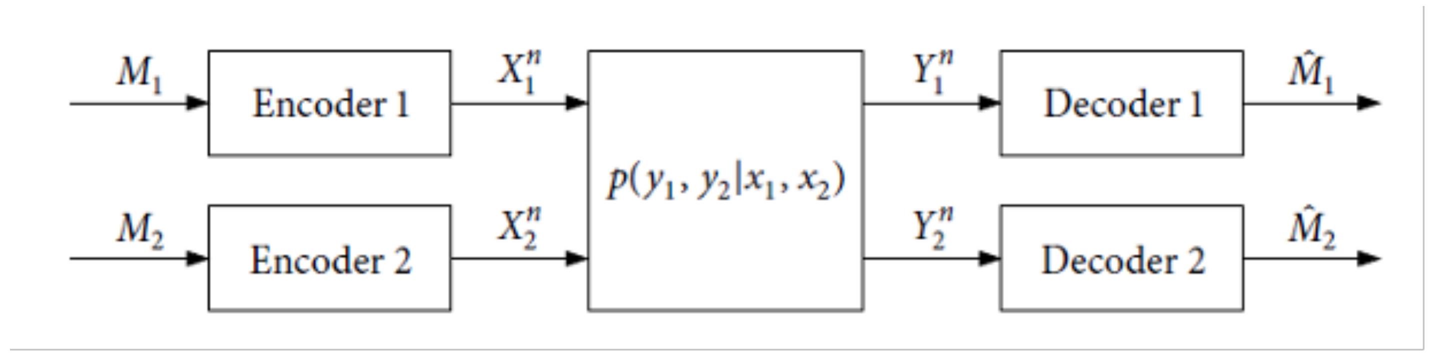
脏纸编码

基本信息论模型

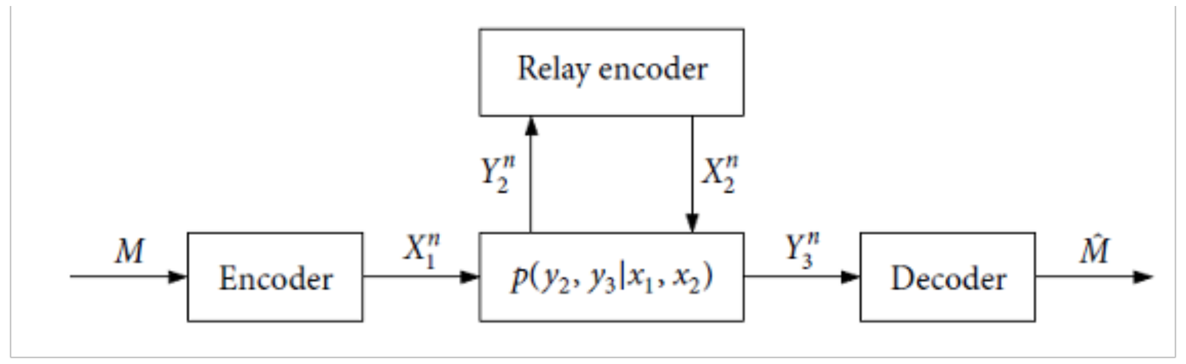
长达40年的未解难题，短期很难突破



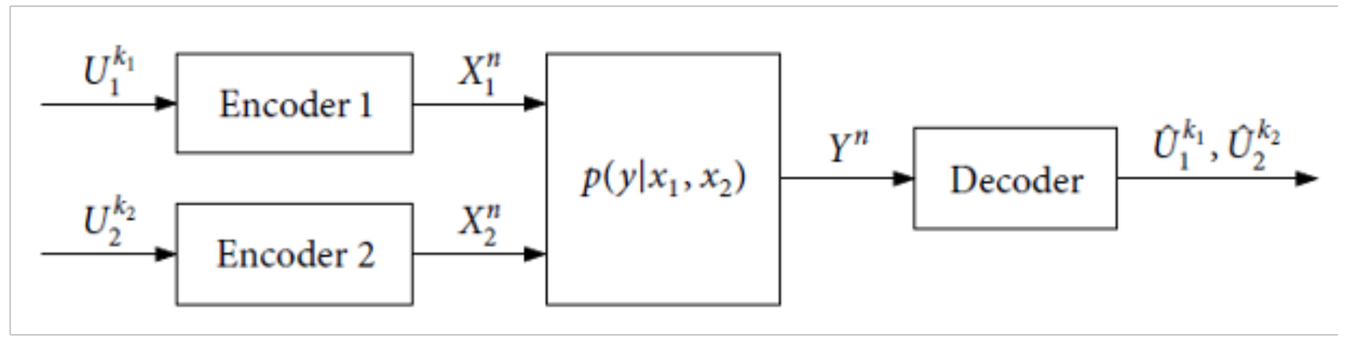
广播信道



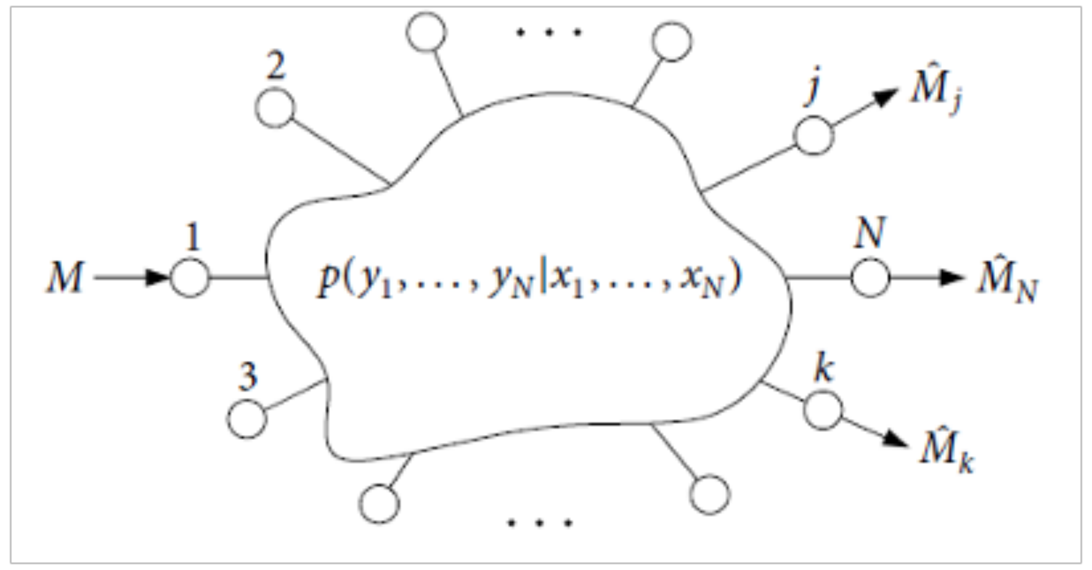
(高斯)干扰信道



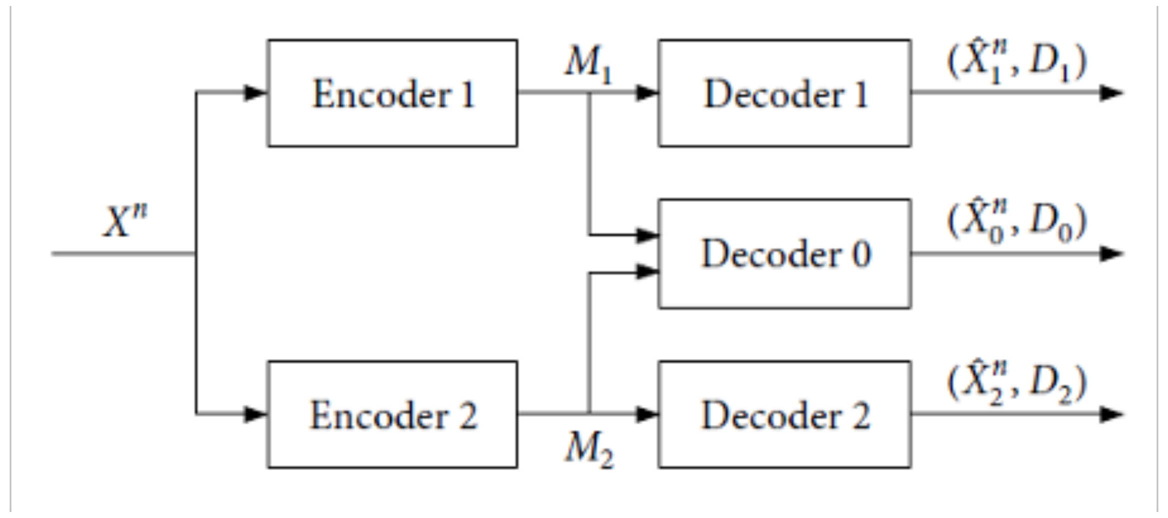
中继信道



多信源信道联合编码



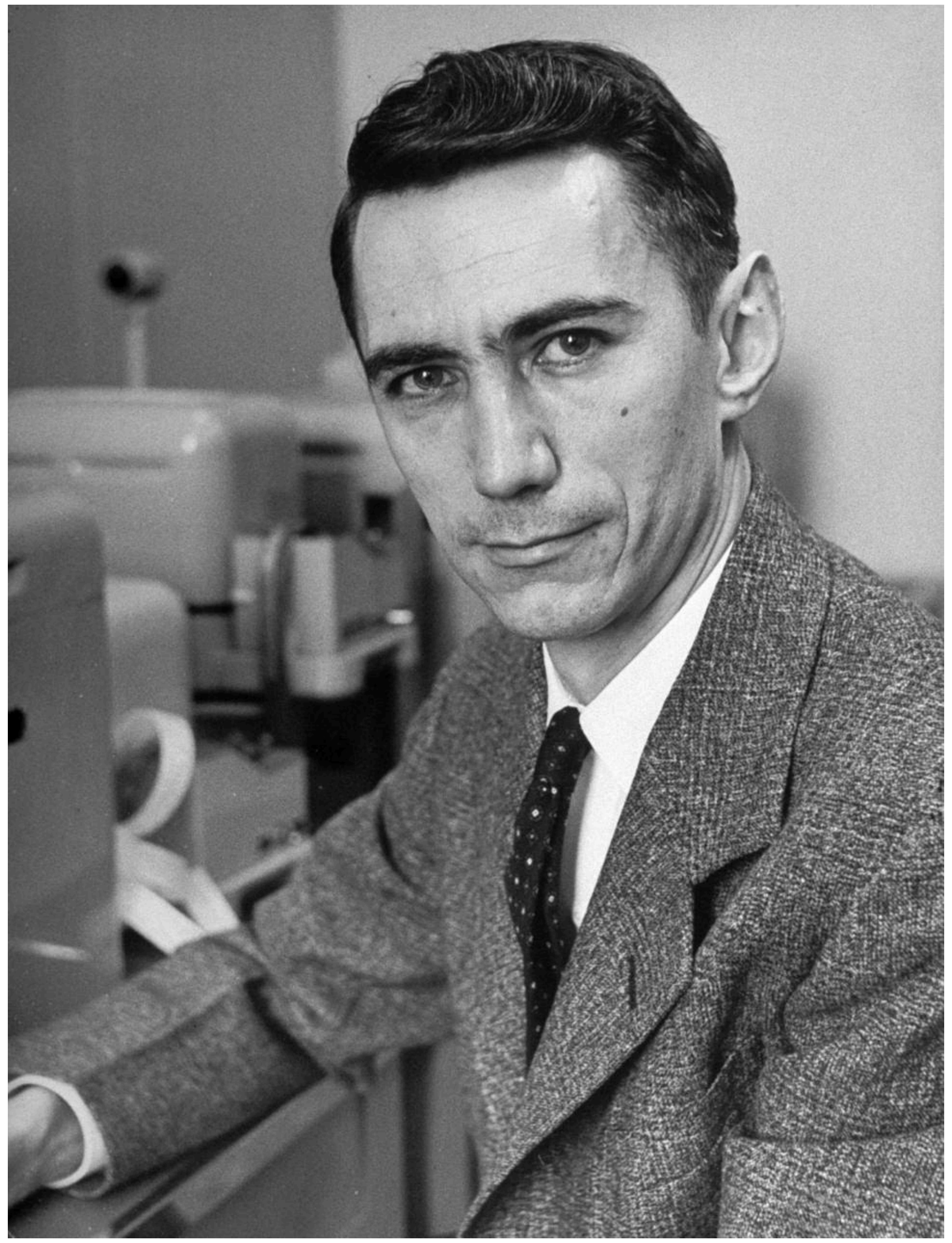
离散无记忆多播网络



多描述编码

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

呼唤新的数学理论



世纪巨人
一直被模仿，很难被超越

C. E. Shannon是信息论之父，他的硕士论文奠定了电路的数学基础。但他却是数学博士，是数学家。

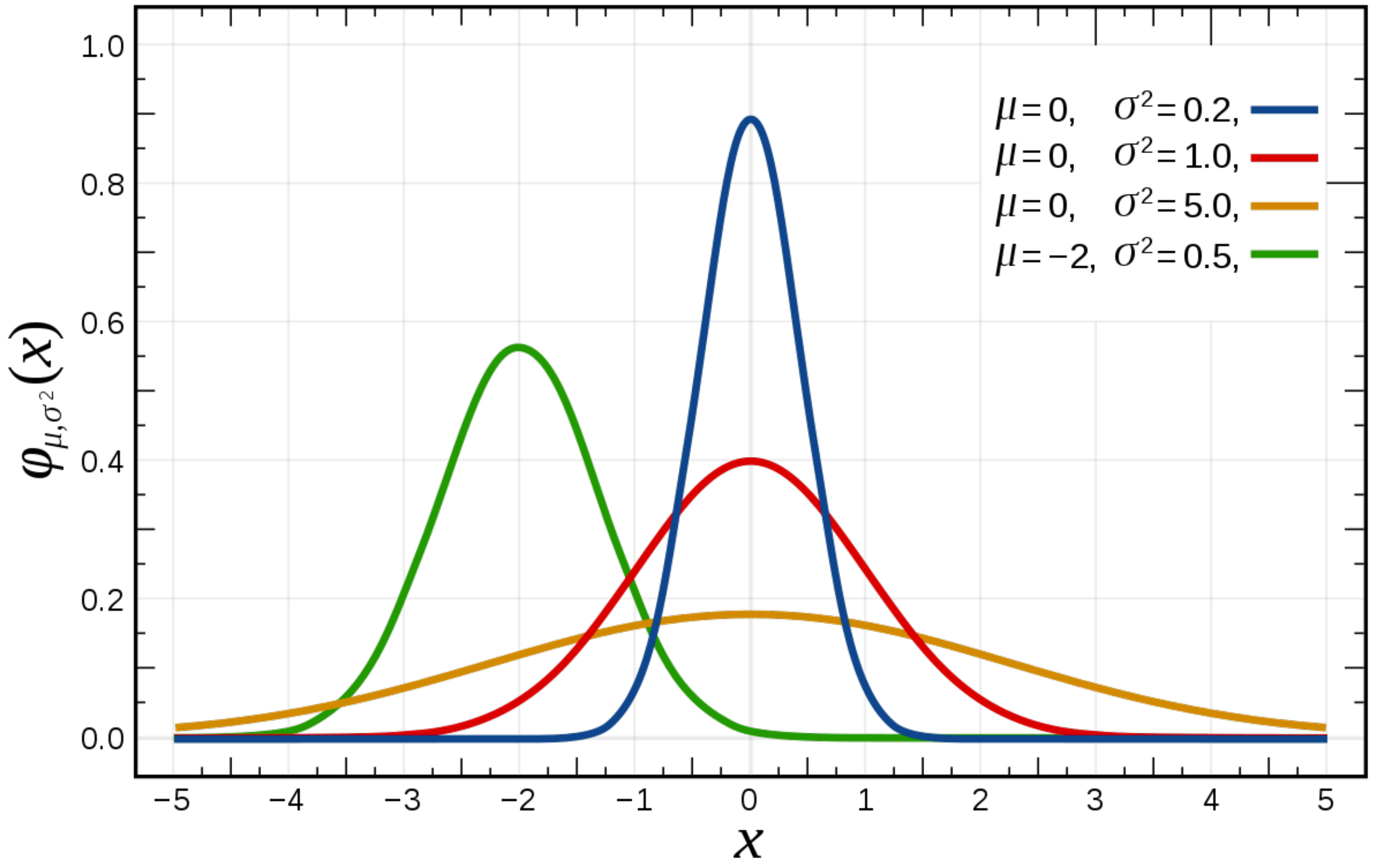
他对那个时代的数学知识有着深刻的理解和把握，使用上的集大成者。

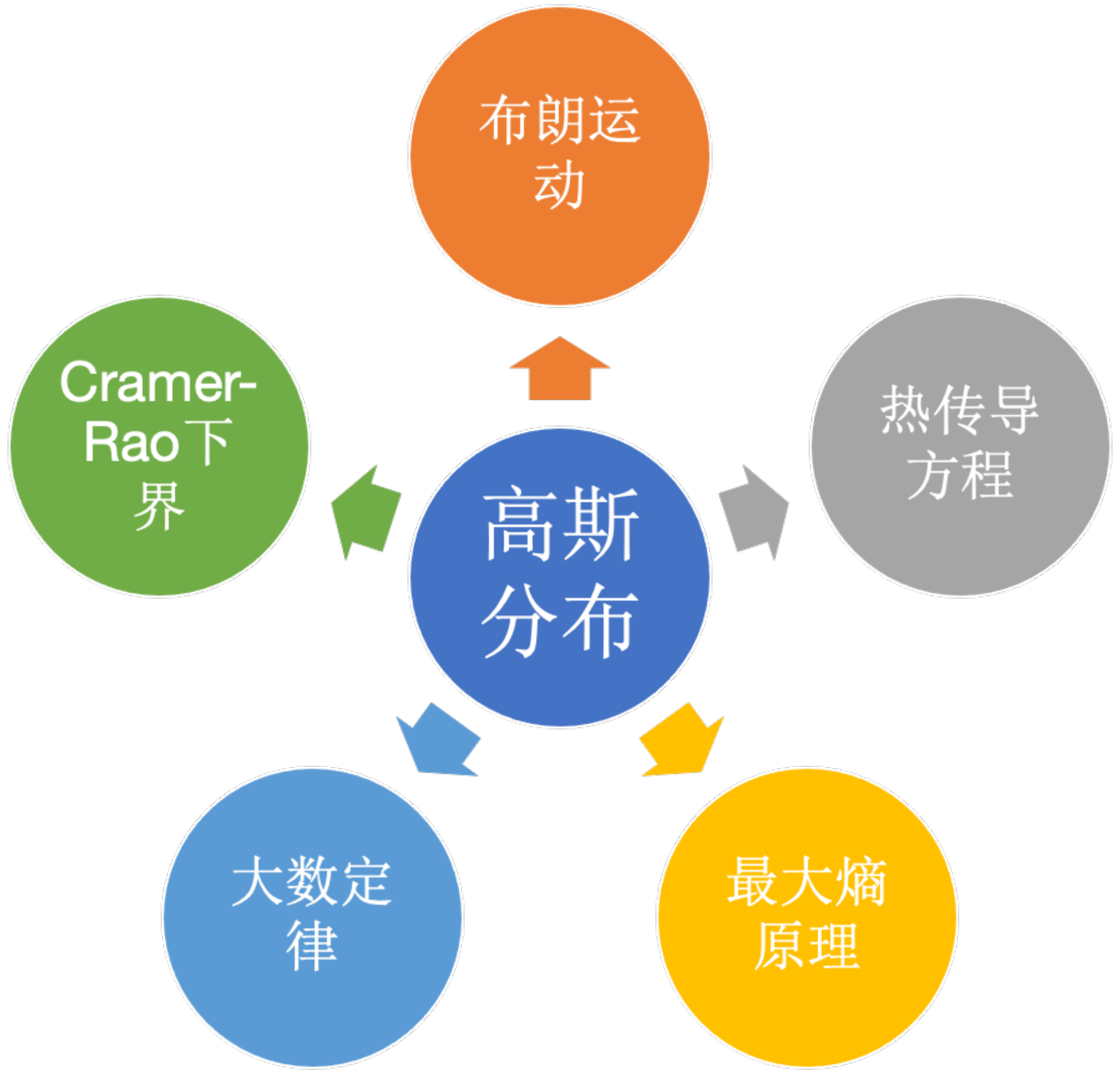
以他独一无二的洞察力和观察力，他的理论也就很难在短期内被超越了。

高斯分布的完全单调猜想

高斯分布

科学与工程的基石





高斯噪声的信息理论

Entropy power inequality (EPI) 熵幂不等式

“高斯分布是最坏的加性噪声”

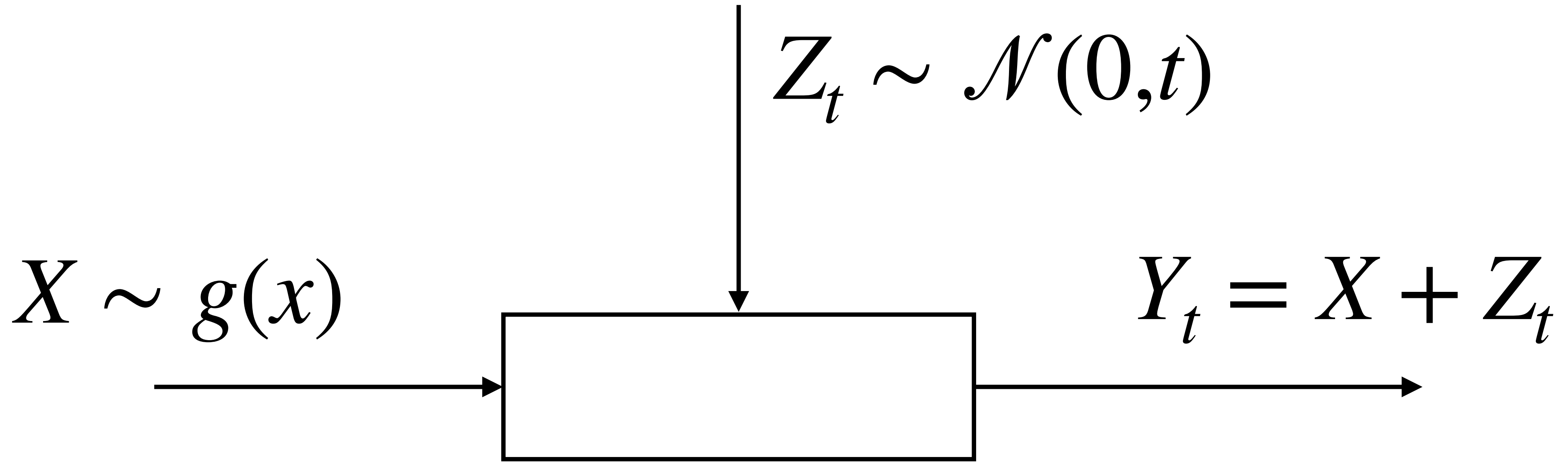
- (Shannon,1948) 对于任何两个独立的连续型随机变量X, Y

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

- 香农理论中最重要的数学工具: 蕴含不确定性原理、等周长定理等等
- 挑战：高斯干扰信道的信道容量一直无法通过EPI获得
    - 一直试图改进，从来没成功
    - 共识：不存在新的更强的EPI，必须开辟新路

高斯信道

和热传导方程等价



- 任意分布的信息X，受到独立高斯噪声Z_t的影响，收到的信息Y_t为X与Z_t的和
  - Y_t 在统计上对应于高斯混合模型 （机器学习）
  - Y_t的概率密度函数$f(y,t)$满足热传导方程(heat equation)

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

信息熵 $h(Y_t) = -\int f(y,t)\log f(y,t)dy$ **基本的信息量**

$h(Y_t)$关于t的级数展开

信息的意义不明

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

- $\frac{\partial}{\partial t} h(Y_t) = \frac{1}{2} I(Y_t) \geq 0$, I 是Fisher信息
- $\frac{\partial^2}{\partial t^2} h(Y_t) \leq 0$
- $\frac{\partial^i}{\partial t^i} h(Y_t)$, $i \geq 3$的情况下性质不明

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

高斯分布的完全单调理论

带符号的第三、四阶显式表达式

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 $t>0$,

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

高斯分布的完全单调理论

相关论文 2013—2022

- F. Cheng, “Generalization of Mrs. Gerber’s Lemma,” Communications in Information and Systems, vol. 14, no. 2, pp. 79-86, 2014 （该文章完成于2011-2012，发表于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

高斯分布的完全单调猜想

2013—2015

总结：$\dfrac{\partial^i}{\partial t^i}h(Y_t)$的前4阶导数的符号+，-，+，-，?，?，?

猜想：后续符号是不是依然保持+，-

- (猜想1) 对于任意分布的随机变量X，$\dfrac{\partial^i}{\partial t^i}h(Y_t)$的符号
    - 在i为偶数的情况，恒为负号
    - 在i为奇数的情况，恒为正号
- (猜想2) $I(Y_t)$ 是log-convex函数

1966年以来的突破

C. Villani教科书中的问题：McKean 1966, 完全单调

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

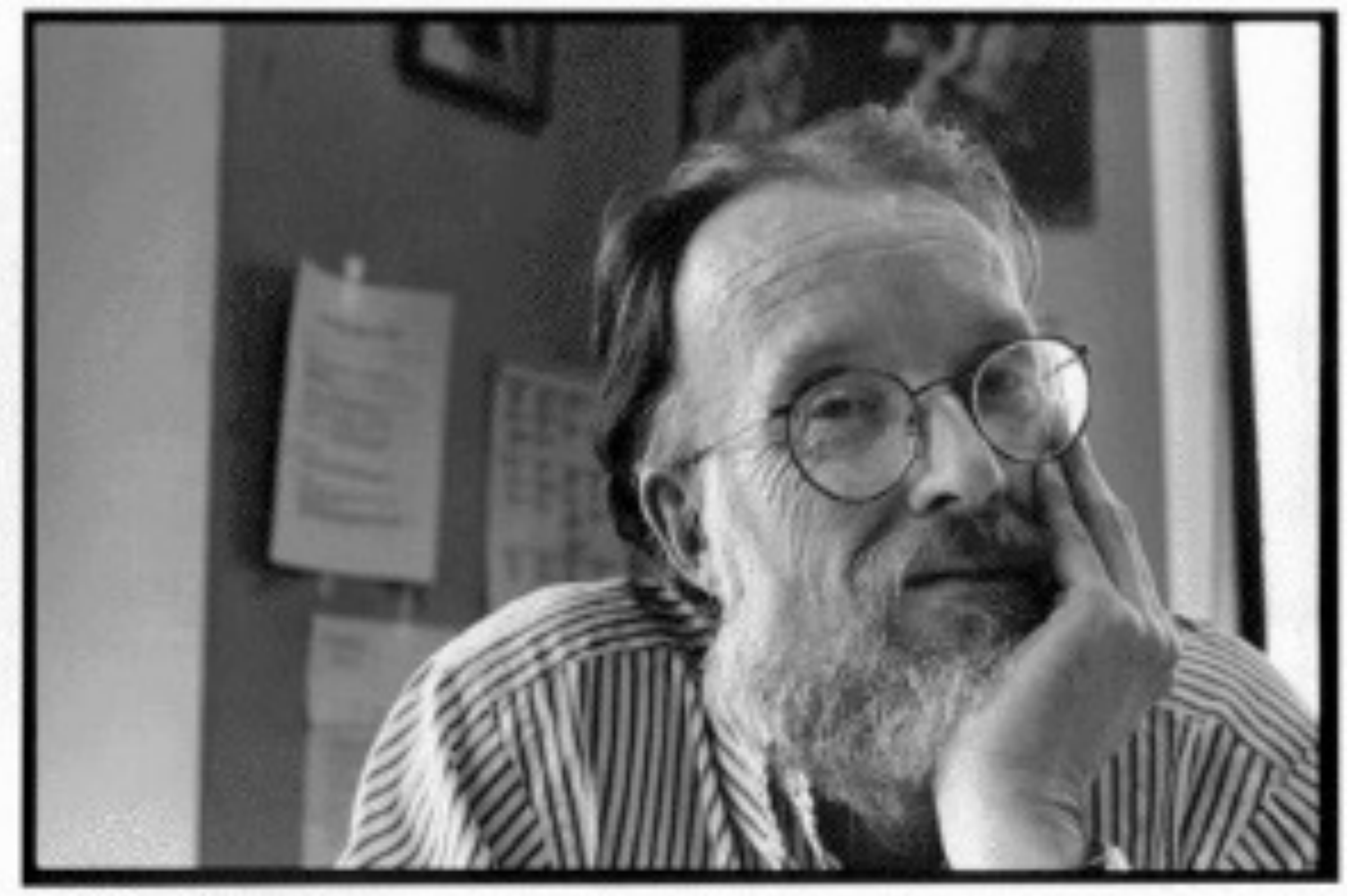
most recent corrections: June 7, 2006

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

完全单调函数理论

Hausdorff—Bernstein—Widder 1920年代

F. Hausdorff

S. N. Bernstein

D. Widder

- 例子：$1/t$的导数符号，-，+，-，+, ……
- 一个函数$f(t)$，如果它的导数的符号正负交替，则称之为完全单调(completely monotone)
  - 求导过程是很机械的，证明符号恒为正或者负一般不容易

高斯分布完全单调猜想（Gaussian Completely Monotone Conjecture, GCMC）

1. Fisher信息$I(Y_t)$ 完全单调
2. $I(Y_t)$ 是log-convex函数

完全单调函数的两个基本性质

H.B.W. 1920

1. $f(t)$完全单调，则$f(t)$必然是log-convex

- 高斯分布完全单调猜想(GCMC)

  1.Fisher信息$I(Y_t)$ 完全单调

  2.$I(Y_t)$ 是log-convex函数

- 猜想1蕴含猜想2

2.Laplace Representation of CM functions

- $f(t)$完全单调，当且仅当存在$[0, +\infty)$上的非递减(non-decreasing)测度$\mu(x)$，使得

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

信息熵的新表达式

这个理论确实很漂亮，但是它的应用是什么?

完全单调猜想的应用以及在干扰信道的验证

完全单调猜想的数学意义

HKU2019： 信息可分解、可逆

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

- Fisher信息的分解
  - e^{-xt}表示完全单调性
  - $\mu(x)$表示X的身份，与t无关
- Laplace变换可逆
  - 已知$Y_t = X + Z_t$: 虽然Z_t干扰了X，但是，$\mu(x)$依然保持不变，依然可逆
  - 回顾，熵增定律，状态不可逆

完全单调猜想的信息论意义

数学性质决定信息意义

$$X \sim g(x) \longrightarrow \boxed{p(Y|X)} \longrightarrow Y \sim f(y)$$

- X经过信道作用，变为Y
- X和Y的相互关系由状态转移矩阵$p(Y|X)$定义
- 迄今为止，没有理论约束X和Y之间的关系
    - 多一个节点，难度天壤之别
- 信息论基础模型一直进展缓慢的技术原因

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

给出了高斯噪声节点的一个数学约束：

可逆、可分解

预测：在高斯干扰信道有应用

高斯多用户信息论中的应用

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

验证：技术进展

$I(Y_t)$是log-convex

- Ledoux等的研究证明了：$I(Y_t)$是log-convex
- 高斯分布的完全单调猜想(GCMC)

  1.Fisher信息$I(Y_t)$ 完全单调

  2.$I(Y_t)$ 是log-convex函数

  - 猜想1蕴含猜想2

想象：猜想2(定理)只是猜想1的一个点!!!

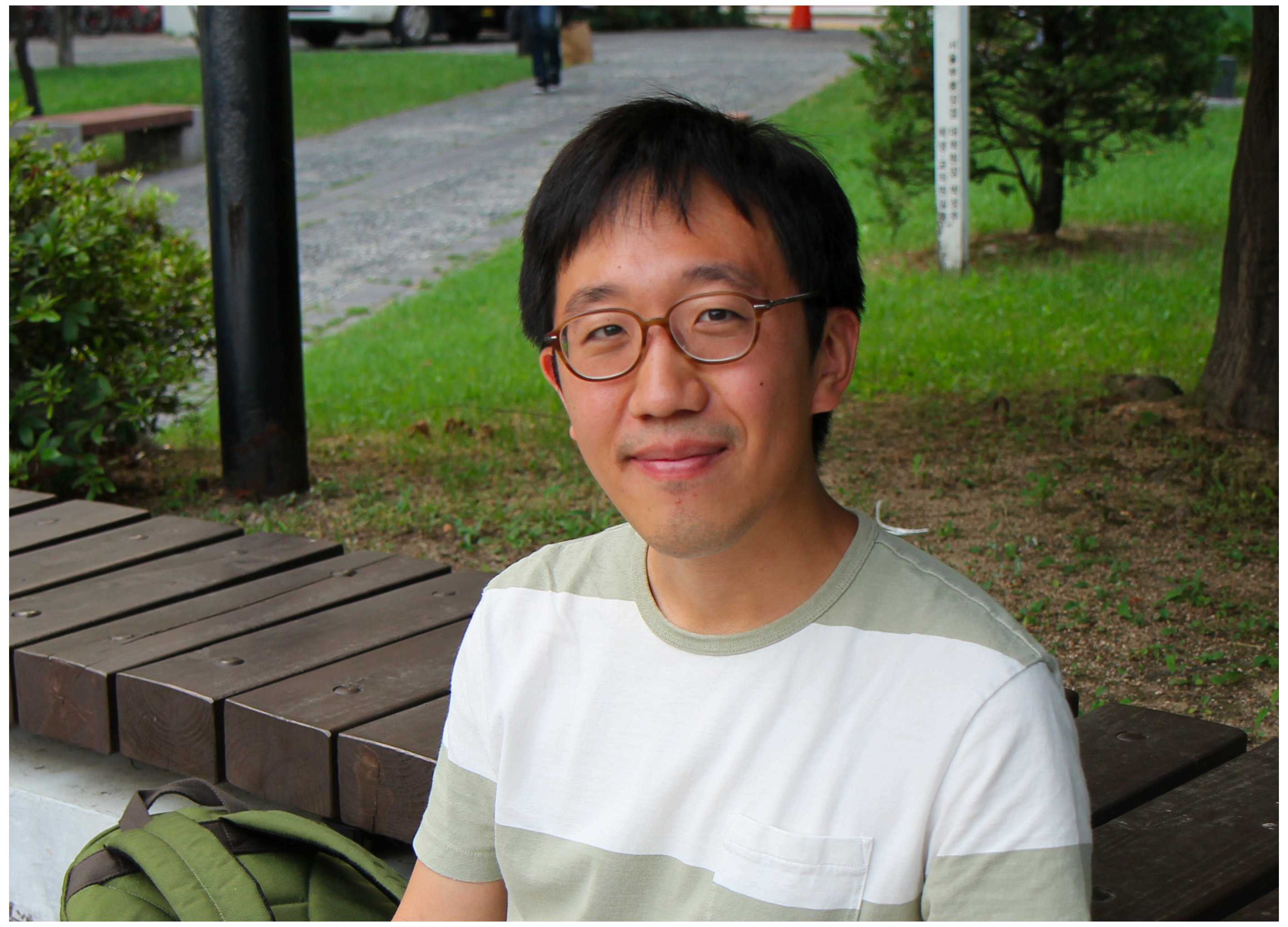
猜想2(定理)：无心插柳，还有更大的惊喜

- 猜想2被验证，给猜想1提供了理论上的证据

完全单调猜想与Hodge理论

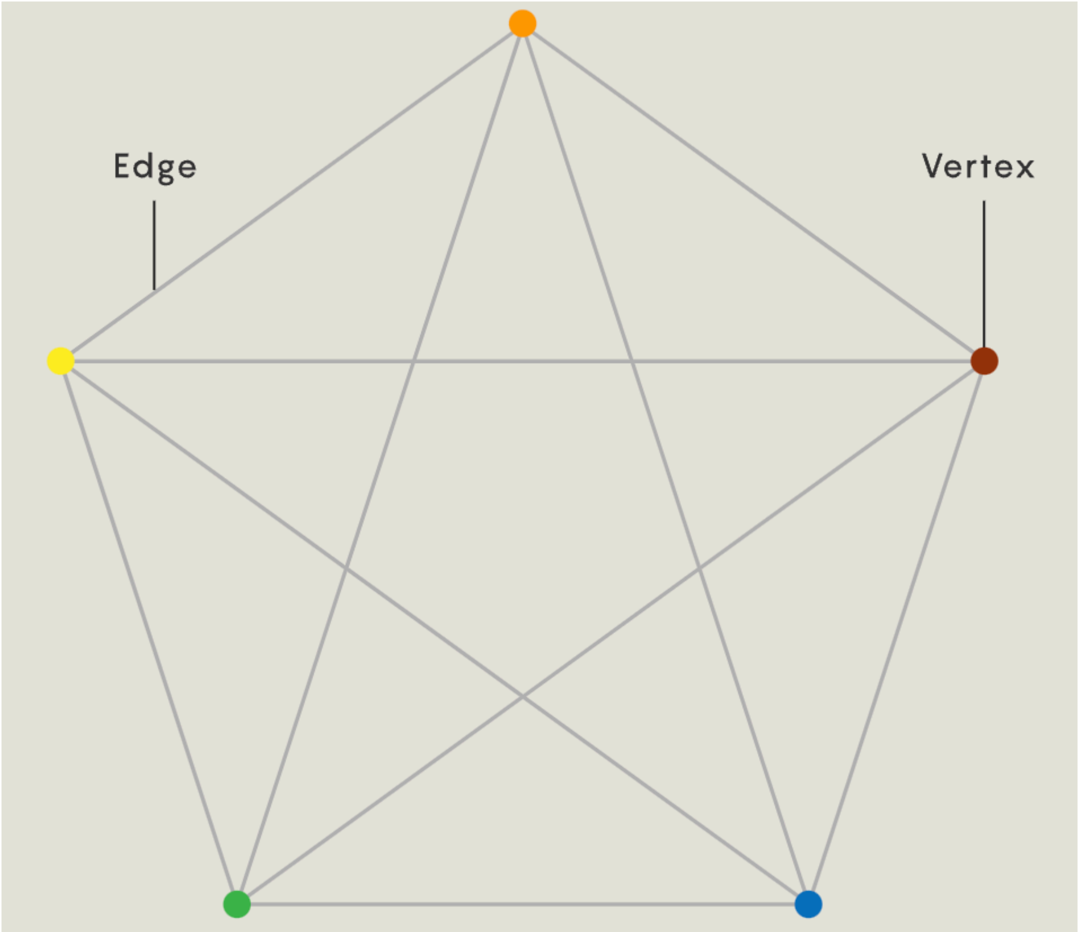
2022国际数学家大会

2022.07.06—2022.07.14



June Huh（许埈珥）
2022 Fields Medal
韩国数学家

- 四色定理（机器证明）
- Chromatic polynomial（染色多项式）
- C: 用n种颜色染色的方案数
  $C = n^5 - 10n^4 + 35n^3 - 50n^2 + 24n$
- C的系数a_i满足 $a_i^2 \geq a_{i-1}a_{i+1}$
- 如果一个序列$\{a_i\}$如果满足
  $a_i^2 \geq a_{i-1}a_{i+1}$，则称为log-concave序列
- **猜想**：所有的C的序列都满足log-concave



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

代数几何构造性地证明log-concave

解决了多个log-concave的开放性问题

- June Huh使用Hodge理论研究组合学
  - 构造Complex algebraic variety, 研究homology和cohomology
- 问题1：chromatic polynomial
- 问题2：matroid
- 问题3：geometry lattice

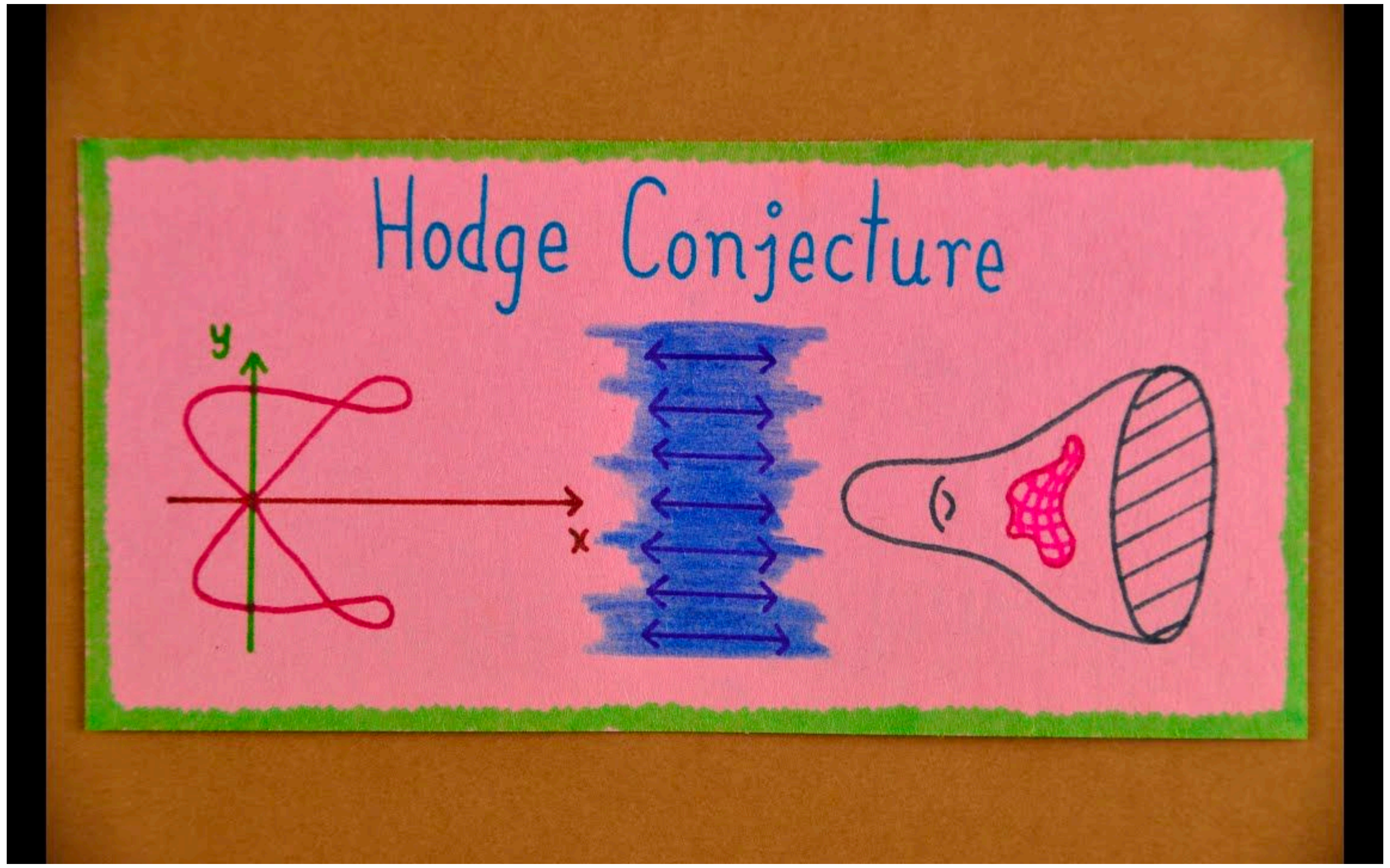
"将霍奇理论的思想引入组合学，证明了几何格的Dowling–Wilson猜想，证明了拟阵的Heron–Rota–Welsh猜想，发展了洛伦兹多项式，以及证明了强梅森猜想。"—Fields奖的评价

Hodge理论

当代数学研究的核心议题

- 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



完全单调函数与Hodge理论

建立和代数几何的联系

- Hausdorff等已证明: 如果 $f(t)$是完全单调函数, 那么$f(t)$是log-convex函数
  - 具体而言: $(\log f(t))'' \geq 0 \rightarrow \dfrac{f''f-(f')^2}{f^2} \geq 0 \rightarrow f''f \geq (f')^2$

    f, f', f'' 是一个log-convex序列
- 完全单调函数的新刻画: 如果 $f(t)$ 是一个完全单调函数, 那么

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

  是一个log-convex序列.
- 如果$\{a_i\}$是一个log-convex序列，那么$\{1/a_i\}$是一个log-concave序列

分析：June Huh的方法→完全单调猜想

- 此地省略2万字

June Huh总结

多篇文章中总结其方法的通用性

> 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

$h(Y_t)$关于t的级数展开

信息论意义

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

- 高斯分布的完全单调猜想(GCMC)

  1.Fisher信息$I(Y_t)$ 完全单调

  2. $I(Y_t)$ 是log-convex函数

$h(Y_t)$表示Y_t的信息熵

熵（bit）可以进一步分解，有内部结构

（例：原子分解为夸克）

- $\{\frac{\partial^i}{\partial t^i} h(Y_t), i = 1,2,\ldots\}$ 后面有一个Hodge结构支撑

- 代数几何名家辈出，浩如烟海，博大精深，Hodge理论只是一个分支
- 高斯完全单调猜想也仅仅是信息理论的一个难题

信息理论的代数几何化

总结与讨论

香农理论的代数几何化

信息理论2.0

- 用代数几何研究编码构造有悠久的历史，理论意义
- 香农理论过去的工具都是分析的，用代数几何研究属于**史无前例**
- Shannon 1948年创建的信息理论主要使用了19世纪的数学
- 代数几何主要是20世纪发展起来的主流数学工具（Grothendieck）
- 代数几何引入到香农理论，将逐步替代Shannon原有的数学基础
- 总结、概括：诞生新一代香农理论（信息论2.0）

信息论2.0

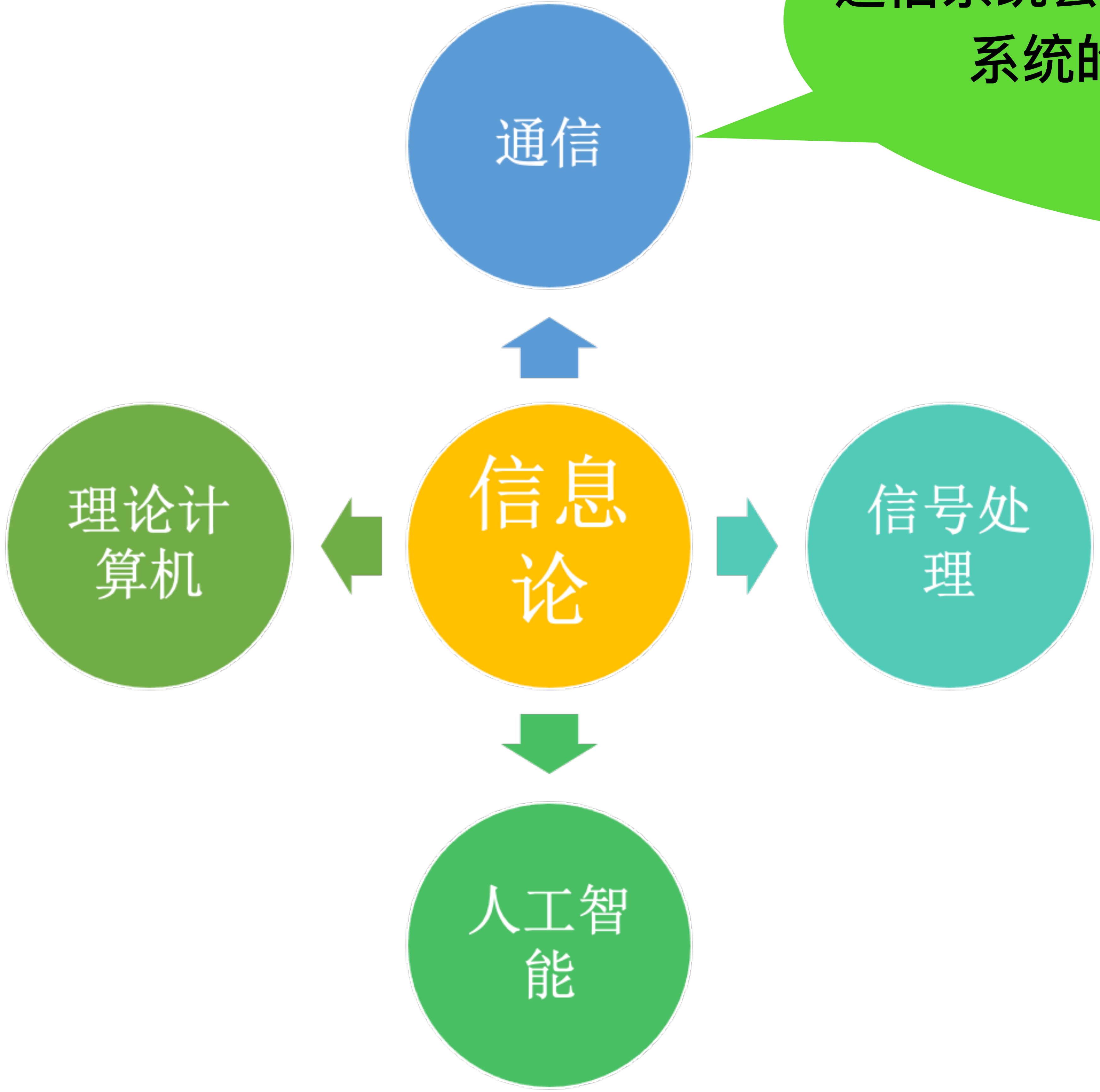
全面代数几何化

- 基于代数几何的数学基础工具
  - 可以汲取最前沿的数学思想（Hodge Conjecture）
- 研究工具的精细：从10^{-3}到10^{-6}乃至10^{-9}
- 预测标志性成果：一些过去很难的信息论问题可能被解决
  - 高斯干扰信道（完全单调猜想的应用）
  - 广播信道
  - ……

对于信息的理解会有一个全新的变革

信息论2.0对信息科学界的影响

数学基础决定上层建筑



通信系统会有全新的研究工具、全新的理解，系统的原理、构造、算法都会变革

我们传递的不再仅仅是bit，而是同调群、Hodge结构！！

技术上的门槛

几代人的代数几何梦想

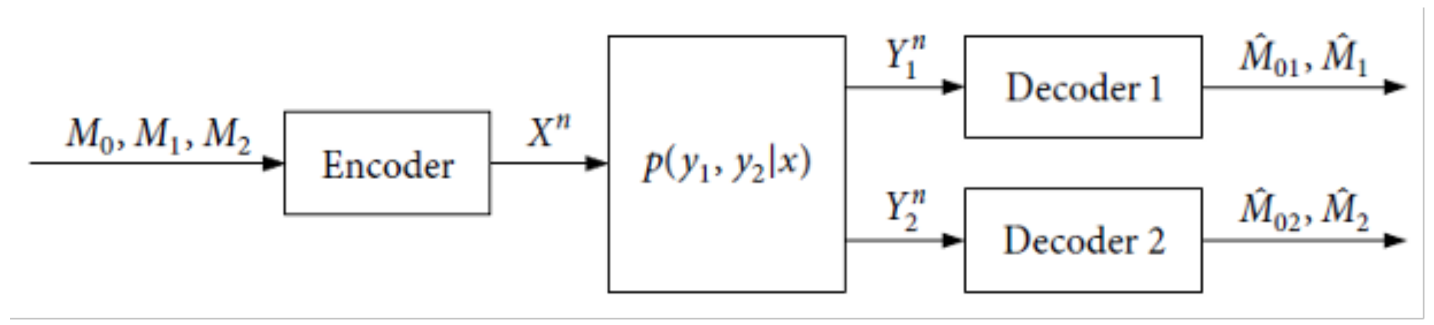
做理论的可能都有代数几何梦想
能否解决实际问题?

- 代数几何门槛太高
  - 入门：近世代数、拓扑、代数拓扑，复分析、交换代数、微分几何、黎曼几何、GTM52
- 世界范围内同时精通代数几何、香农理论的人基本没有
- 需要香农理论、代数几何、复分析、代数拓扑、代数组合、编码理论紧密合作
- 通信产业的人员跟进（实现、算法、专利）

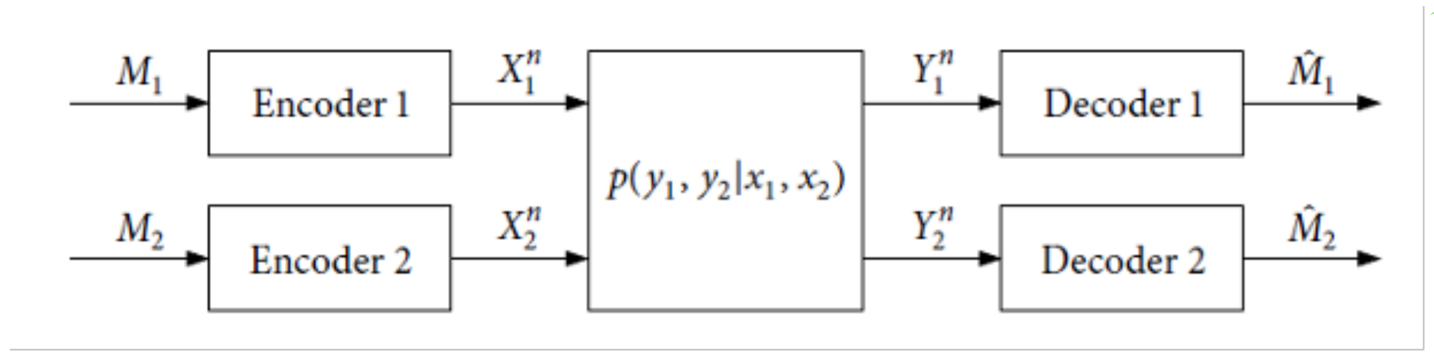
信息科学的难题有多难

可能和代数几何中的著名难题(Hodge理论)有关

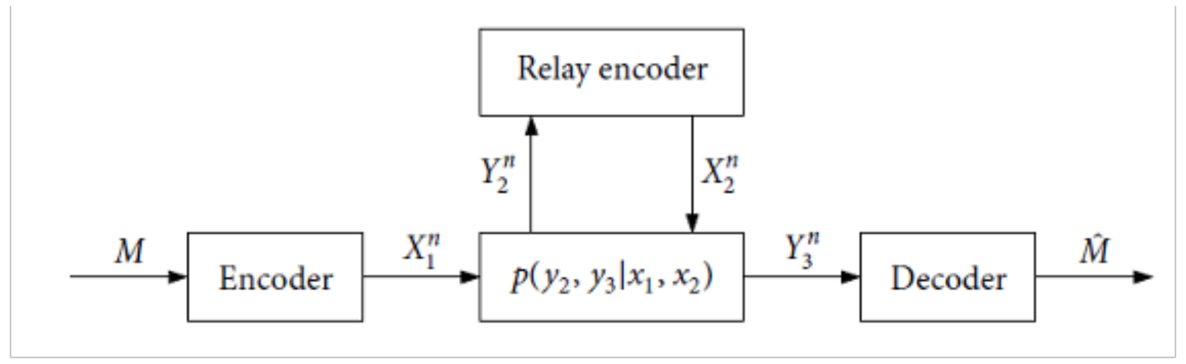
其实我们一直在努力 :)



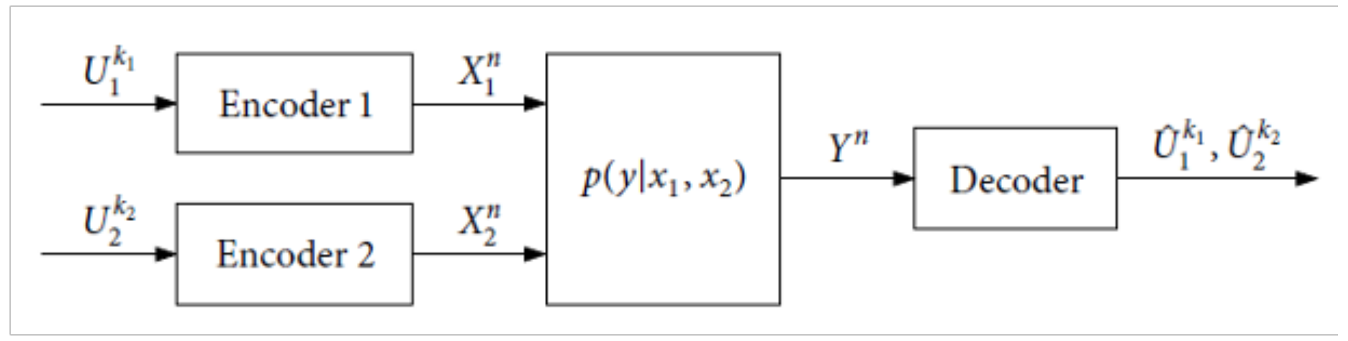
广播信道



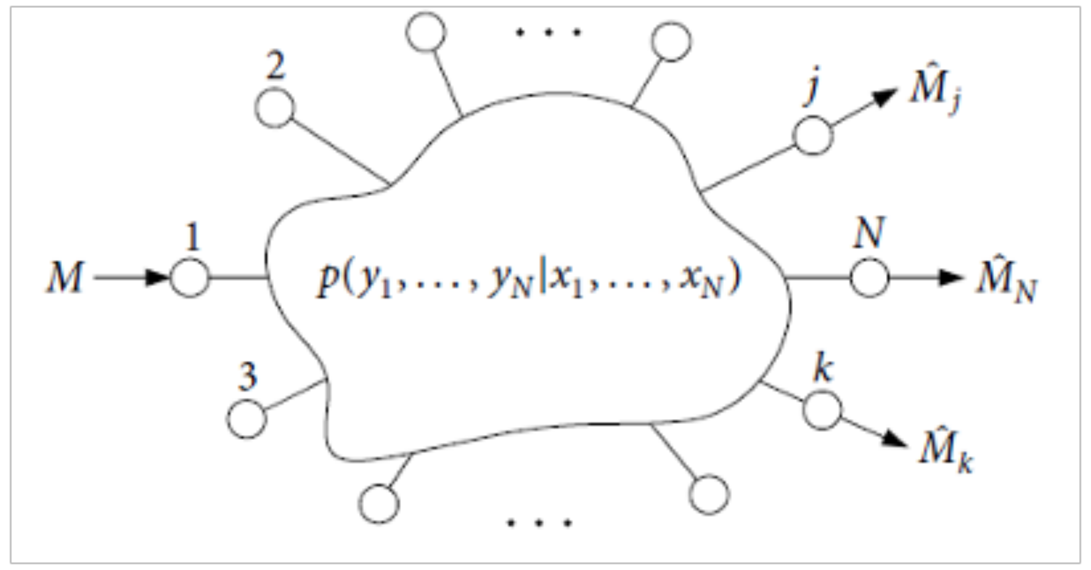
(高斯)干扰信道



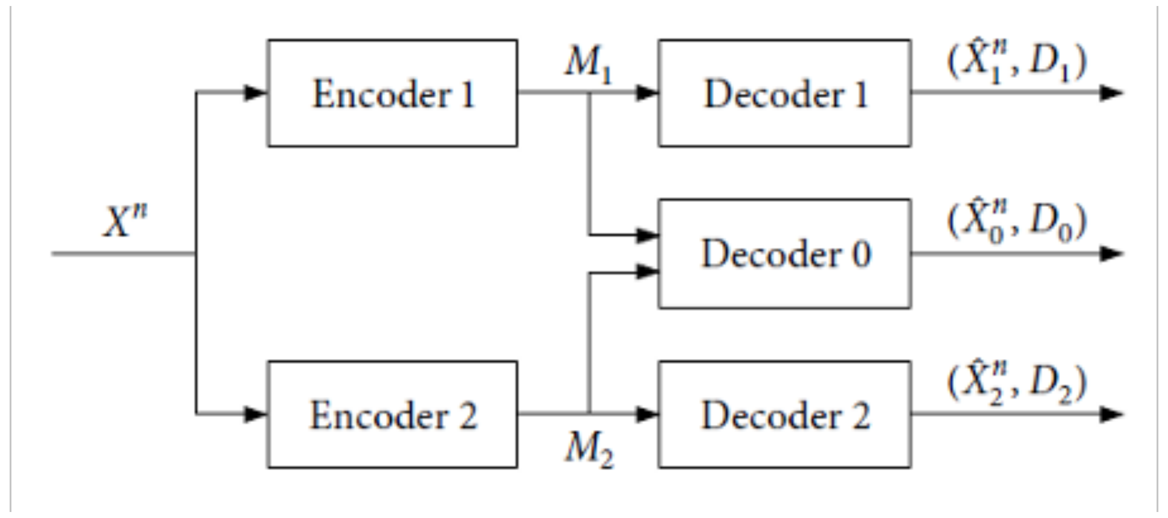
中继信道



多信源信道联合编码



离散无记忆多播网络



多描述编码

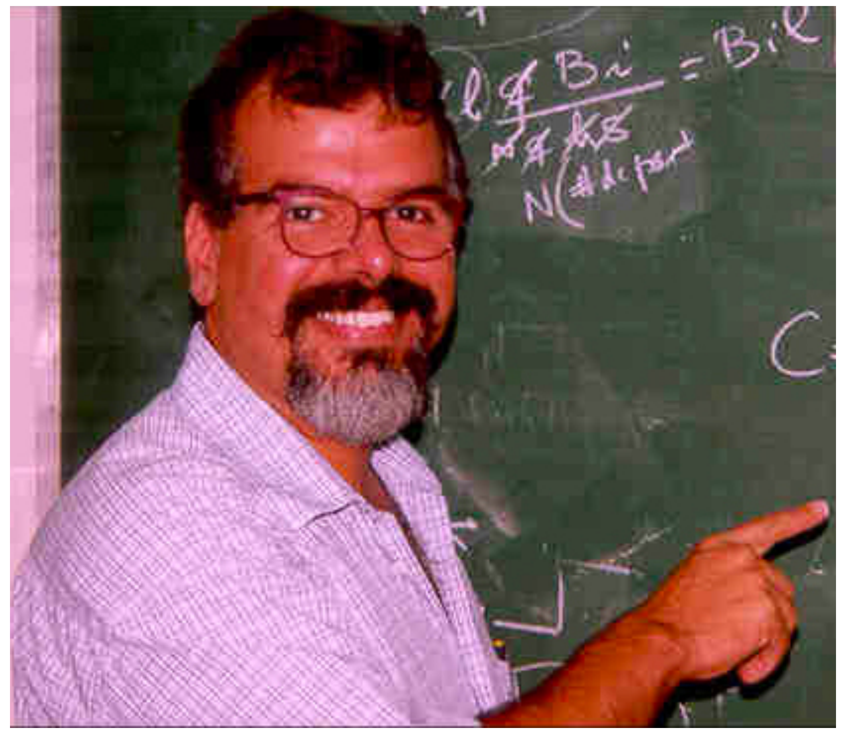
信息论的代数几何化（2.0）已经开始

近3年会有一些根本性改变


C. E. Shannon


T. Cover

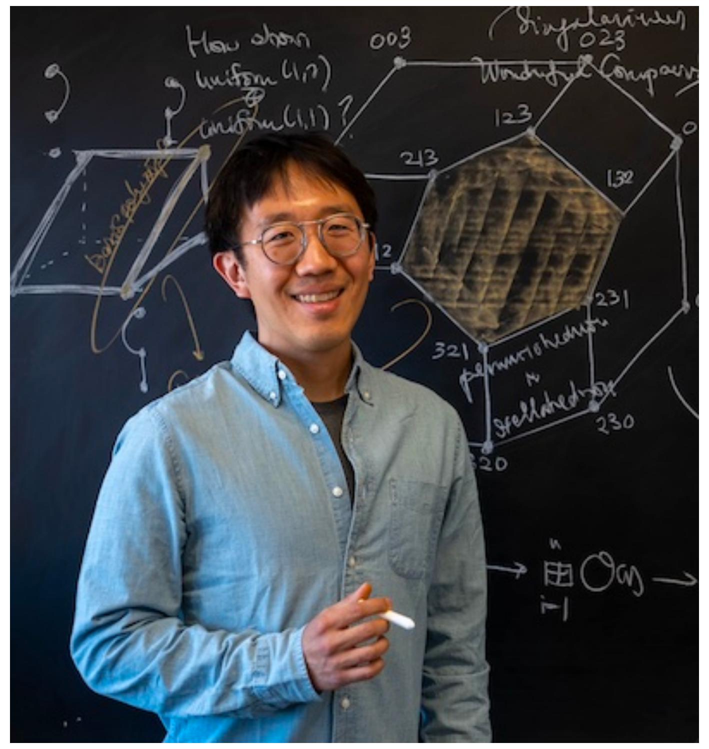

A. El Gamal


M. H. Costa


R. W. Yeung


F. Hausdorff


C. Villani


J. Huh

一切过往，皆为序曲
基础研究的意义

谢谢大力支持

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/