Chaos, Solitons and Fractals 184 (2024) 114976

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

Generalized information entropy and generalized information dimension

Tianxiang Zhan [a], Jiefeng Zhou [a,b], Zhen Li [c], Yong Deng [a,d,*]

[a] Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, Sichuan, China
[b] Yingcai Honors College, University of Electronic Science and Technology of China, Chengdu, 610054, Sichuan, China
[c] China Mobile Information Technology Center, China Mobile, 100029, Beijing, China
[d] School of Medicine, Vanderbilt University, Nashville, 37240, TN, USA

ARTICLE INFO

Keywords:
Information entropy
Probability
Mass function
Approximate calculation
Coding theory

ABSTRACT

The concept of entropy has played a significant role in thermodynamics and information theory, and is also a current research hotspot. Information entropy, as a measure of information, has many different forms, such as Shannon entropy and Deng entropy, but there is no unified interpretation of information from a measurement perspective. To address this issue, this article proposes Generalized Information Entropy (GIE) that unifies entropies based on mass function. Meanwhile, GIE establishes the relationship between entropy, fractal dimension, and number of events. Therefore, Generalized Information Dimension (GID) has been proposed, which extends the definition of information dimension from probability to mass fusion. GIE plays a role in approximation calculation and coding systems. In the application of coding, information from the perspective of GIE exhibits a certain degree of particle nature that the same event can have different representational states, similar to the number of microscopic states in Boltzmann entropy.

1. Introduction

Entropy, as a physical quantity, plays a role in many fields and is one of the research hotspots in many disciplines [1–3]. The concept of entropy dates back to thermodynamics and was first named by Clausius [4]. Boltzmann proposed Boltzmann entropy formula which is a view of Statistical Physics [5]. Furthermore, Gibbs proposed the Gibbs entropy based on probability measure [6]. Then, in information theory, Shannon entropy is proposed and also is based on probability measure [7]. Shannon entropy is widely applied in many areas such as distance [8,9], coding [10,11], statistics [12,13] and so on.

There are many different forms of information entropy. In terms of information discreteness and continuity, the information entropy of continuity has also been explored. Differential entropy is a continuous information entropy [14]. The dimension of differential entropy is problematic, so Jaynes proposed the Jaynes entropy for correction [15–17]. Jaynes entropy plays an important role in continuous information, especially in magnetism [18,19]. On the other hand, information can not only be represented by probability. Probability is a special case of mass function [20,21]. Then, Deng entropy was proposed to measure the information represented by mass functions [22]. The application of Deng entropy is extensive [23–25]. Considering there is not only combination information but permutation information, Random Permutation Set (RPS) and its entropy is proposed to describable the order of samples [26]. The forms of different entropy have not been effectively unified, and fundamentally different entropy measures information from different perspectives. A unified form of entropy is beneficial for exploring the essence of information.

Another significant concept worthy of mention is the information dimension. The fractal dimension, which quantifies the capacity of fractal expansion, serves as the precursor to the information dimension, initially introduced by Rényi [27]. The information dimension, being closely associated with information entropy, serves as an indicator of the rate of information accumulation. This metric finds widespread application in various domains, including the analysis of stochastic processes [28], data compression techniques [29], and inference algorithms [30], among others. The information dimension of Deng entropy and RPS entropy has also been proposed [31,32]. The information dimension can be used to calculate information entropy, which was proposed by Rényi, but it has not been extended to entropy based on mass function. Therefore, an information dimension corresponding to a unified form of entropy is also needed.

In order to solve the inconsistency in the form of information entropy, Generalized Information Entropy (GIE) is proposed. GIE is a method of measuring information based on mass function. When the sample space and event space of the mass function are constrained, GIE can degenerate to the previously mentioned entropy, such as Shannon

* Corresponding author at: Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, Sichuan, China.

E-mail addresses: zhen.li@pku.edu.cn (Z. Li), dengentropy@uestc.edu.cn (Y. Deng).

https://doi.org/10.1016/j.chaos.2024.114976
Received 27 January 2024; Received in revised form 22 March 2024; Accepted 6 May 2024
Available online 14 May 2024

entropy, Deng entropy. On the other hand, corresponding to GIE, the Generalized Information Dimension (GID) has also been proposed. GID also describes the growth rate of information and is also a fractal dimension. Based on GID, an approximate formula Approximate Generalized Information Entropy (AGIE) is proposed to establish the quantitative relationship between information entropy and information dimension.

GIE and GID have many application scenarios. A common problem faced by information entropy is that there may be exponential explosions in the event space, leading to the incalculability of information entropy. The proposed AGIE effectively solves this problem, providing a method for estimating GIE, and as the sample space increases, the estimated error converges to 0. Another application scenario is encoding. In the perspective of GIE, the same event can represent different states, similar to particles.

The structure of the article is as follows. In Section 2, the relevant preliminaries are introduced. In Section 3, GIE and GID are introduced. In Section 4, it shows some properties of GIE and GID. In Section 5, the applications of GIE and GID is shown. In Section 6, the article summarizes the full text and looks forward to future research topic.

2. Preliminary

2.1. Boltzmann entropy

Boltzmann entropy H_b is defined as Eq. (1) [5].

$$H_b = k_b * ln(W) \tag{1}$$

where $k_b \approx 1.380649 * 10^{-23}$ J/K is Boltzmann constant which is able to regard as increasing coefficient and W is the number of microstate of the system.

2.2. Shannon entropy

Shannon Entropy H_s is defined as Eqs. (2) and (3) [7]. Ω is discrete sample space containing exclusive and exhaustive samples θ_i as Eq. (4) and $n = |\Omega|$ is size of Ω.

$$H_s = -\sum_{\theta_i \in \Omega} p(\theta_i) * log_2(p(\theta_i)) \tag{2}$$

$$\sum_{\theta_i \in \Omega} p(\theta_i) = 1 \tag{3}$$

$$\Omega = \{\theta_1, \theta_2, \theta_3, \dots, \theta_n\} \tag{4}$$

2.3. Jaynes entropy

Jaynes entropy is continuous version of Shannon Entropy and is defined in Eq. (5) where $m(\theta)$ reflects invariant measure [15–17].

$$H_j = \int_{\Omega} p(\theta) log_2(\frac{p(\theta)}{m(\theta)}) d\theta \tag{5}$$

2.4. Deng entropy

Evidence theory is an extension of probability theory [20,33], where mass function is assigned to power set. Power set $\mathcal{E}_c$ of Ω in Eq. (4) is as Eq. (6). Mass function of $\mathcal{E}_c$ is defined in Eq. (7).

$$\mathcal{E}_c = \{\{\theta_1\}, \{\theta_2\}, \dots, \{\theta_n\}, \{\theta_1, \theta_2\}, \dots, \Omega\} \tag{6}$$

$$\sum_{i \in \mathcal{E}_c} m(i) = 1 \tag{7}$$

Given a group of mass function m in Eq. (7), its corresponding entropy, named as Deng entropy, is defined in Eq. (8) where $|i|$ is size of event i [22].

$$H_d = -\sum_{i \in \mathcal{E}_c} m(i) log_2(\frac{m(i)}{2^{|i|} - 1}) \tag{8}$$

2.5. Entropy of random power set

According to Pascal triangle [34], the power set can be seen as all possible event combinations, since $2^n - 1 = \sum_{k=1}^{n} \binom{n}{r}$ where $\binom{n}{r}$ is combinatorial number. Similarly, all possible event permutations can be defined as Random Permutation Set (RPS) in Eq. (9) [26].

$$\begin{aligned} \mathcal{E}(\Omega) &= \{A_{ij} | i = 0, \dots, N; j = 1, \dots, P(N, i)\} \\ &= \{(\theta_1), (\theta_2), \dots, (\theta_N), (\theta_1, \theta_2), (\theta_2, \theta_1), \dots, \\ &\quad (\theta_1, \theta_2, \dots, \theta_N), \dots, (\theta_N, \theta_N - 1, \dots, \theta_1)\} \end{aligned} \tag{9}$$

Given a group of permutation mass function $\mathscr{M}$ in Eq. (10), its corresponding RPS entropy is defined in Eq. (11), where $F(i)$ is shown in Eq. (12) where $P(n, k)$ is permutation number.

$$\sum_{i=1}^{N} \sum_{j=1}^{P(N,i)} \mathscr{M}(A_{ij}) = 1 \tag{10}$$

$$H_r = -\sum_{i=1}^{N} \sum_{j=1}^{P(N,i)} \mathscr{M}(A_{ij}) log_2(\frac{\mathscr{M}(A_{ij})}{F(i) - 1}) \tag{11}$$

$$F(i) = \sum_{i}^{j=1} P(i, j) \tag{12}$$

2.6. Fractal dimension

The fractal dimension D is a measure of how a pattern fills space and defined in Eq. (13) where N is number of units and ϵ is scaling factor [35–37].

$$N = \epsilon^{-D} \tag{13}$$

2.7. Rényi dimension

Rényi proposed an another form of Shannon Entropy with Rényi dimension d as Eq. (14) where ξ is a random variable and ξ_n is defined in Eq. (15) [27,29]. In Eq. (14), $o(1)$ represents a remainder term tending to 0 for $n \to +\infty$.

$$H_{\mathcal{E}}(\xi_\eta) = d * log_2 n + h + o(1) \tag{14}$$

$$\xi_n = \frac{1}{n} \lfloor n\xi \rfloor \tag{15}$$

Rényi Dimension is defined in Eq. (16) which is a measure of the fractal dimension of a probability distribution.

$$d(\xi) = \lim_{n \to +\infty} \frac{H_{\mathcal{E}}(\xi_n)}{log_2 n} \tag{16}$$

3. Proposed entropy and dimension

3.1. From probability to mass function

A Probability Space $(\Omega, \mathcal{F}, p)$ is a triple where Ω is sample space, $\mathcal{F}$ is σ-algebra which is also called event space and p is probability measure $p : \mathcal{F} \to [0, 1]$. For $A_i \in \mathcal{F}$, p is countably additive in Eq. (17) and needs to satisfy Eq. (18) . Probability $p(A)$ means that the possibility of event A happening.

$$p(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} p(A_i) \tag{17}$$

$$\sum_{A_i \in \mathcal{F}} p(A_i) = 1 \tag{18}$$

Mass function is a generalized probability [33] shown in Fig. 1. Mass function space $(\Omega, \mathcal{E}, m)$ is also a triple and event space $\mathcal{E}$ is not σ-algebra. Events in Event Space $\mathcal{E}$ is composed of relations of samples in the sample space Ω. Simultaneously, m is a map $m : \mathcal{E} \to [0, 1]$ and m is not countably additive. m needs to satisfy Eq. (19). Mass $m(A)$ means that the confidence of subset of event A happening which is reflected

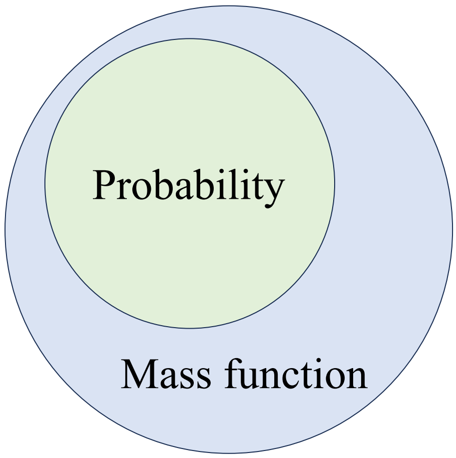


Fig. 1. Relation between probability and mass function.

by encoding in Section 5.4. For example, given two exclusive samples A, B, $m(AB)$ means confidence of A, B or AB happening.

$$\sum_{\theta_i \in \mathcal{E}} m(i) = 1 \tag{19}$$

For mass function, its belief function (Bel) and plausibility function (Pl) are defined in Eqs. (20) and (21) and are upper and lower bounds of probability in Eq. (22) [20,38]. When event space is only composed of events which contain single sample A, the upper bound $Pl(A)$ and lower bound $Bel(A)$ equal to mass function $m(A)$, then mass function is degenerated to probability in Eq. (23).

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{20}$$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{21}$$

$$Bel(A) \leqslant p(A) \leqslant Pl(A) \tag{22}$$

$$p(A) = m(A) \ (\because \ Bel(A) = m(A), \ \ Pl(A) = m(A)) \tag{23}$$

Example 1 (*The Division of Possibilities by Mass and Probability*)**.** According to Eqs. (18) and (19), it shows that mass and probability are essentially partition possibility into different event spaces $\mathcal{F}, \mathcal{E}$. Event space $\mathcal{F}$ of probability only contains events composed of single sample and event space $\mathcal{E}$ of mass contains events composed of combinations. Assuming the sample space is $\Omega = \{A, B\}$, Fig. 2 represents different divisions of mass and probability which correspond to Eqs. (24) and (25).

$$p(A) + p(B) = p(AB) = p(\Omega) = 1 \tag{24}$$

$$m(A) + m(B) + m(AB) = 1 \tag{25}$$

It is noteworthy that $\mathcal{E}$ in mass function space $(\Omega, \mathcal{E}, m)$ is generalized. Event space cannot be combination but permutation even complex information structure shown in Fig. 3. In the event space, events represent the basic relationships between samples. A combination indicates that all samples within the combination are possible to happen, with no priority among the samples, whereas permutation implies a certain level of priority information. Therefore, the generalized information entropy mentioned later only considers the discrete event space and is not compatible with the continuous Jaynes entropy.

3.2. Generalized information entropy

Generalized Information Entropy (GIE) H_g is defined as Eqs. (26) and (27) where Ω is sample space, $\mathcal{E}$ is event space and m is mass function. $\mathcal{E}(\cdot)$ is event space generated from corresponding event.

Definition 1 (*Generalized Information Entropy*)**.**

$$H_g = -\sum_{i \in \mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \tag{26}$$

$$\sum_{i \in \mathcal{E}(\Omega)} m(i) = 1 \tag{27}$$

Example 2 (*Local Similarity*)**.** GIE has fractal characteristics and exhibits local similarity. In the calculation process of GIE, it is necessary to calculate the corresponding event space $\mathcal{E}(A)$ of each event A in event space $\mathcal{E}$ such as an example in Fig. 4. This is a process of local similarity, in which the event space is regenerated from each event.

For a sample space Ω in Eq. (28), its events spaces in form of combination $\mathcal{E}_c$ and permutation $\mathcal{E}_p$ are in Eqs. (29) and (30). For the combination event AB, the generated event space is in Eq. (31); for the permutation event AB, the generated event space is in Eq. (32). These event spaces generated by events have physical significance, representing all possible relationships that can be represented by the event in the encoding.

$$\Omega = \{A, B, C\} \tag{28}$$

$$\mathcal{E}_c = \{A, B, C, AB, AC, BC, ABC\} \tag{29}$$

$$\begin{aligned} \mathcal{E}_p = \{&A, B, C, AB, AC, BA, BC, CA, CB, ABC, ACB, \\ &BAC, BCA, CAB, CBA\} \end{aligned} \tag{30}$$

$$\mathcal{E}_c(AB) = \{A, B, AB\} \tag{31}$$

$$\mathcal{E}_p(AB) = \{A, B, AB, BA\} \tag{32}$$

Example 3 (*Degeneration from GIE Shannon Entropy*)**.** When each event in event space is composed of single sample and mass function equals probability, the GIE is degenerated to Shannon Entropy in Eq. (33).

$$\begin{aligned} H_g &= -\sum_{i \in \mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \\ &\xlongequal{|\mathcal{E}(i)|=1} -\sum_{i \in \Omega} p(i) * log_2(p(i)) = H_s \end{aligned} \tag{33}$$

Example 4 (*Degeneration from GIE to Deng Entropy*)**.** When event space is power set of sample space, the GIE is degenerated to Deng Entropy in Eq. (34).

$$\begin{aligned} H_g &= -\sum_{i \in \mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \\ &\xlongequal{|\mathcal{E}(i)|=2^{|i|}-1} -\sum_{i \in \mathcal{E}(\Omega)} m(i) log_2(\frac{m(i)}{2^{|i|}-1}) = H_d \end{aligned} \tag{34}$$

Example 5 (*Degeneration from GIE to RPS Entropy*)**.** When event space is permutation event space, the GIE is degenerated to RPS Entropy in Eq. (35).

$$\begin{aligned} H_g &= -\sum_{i \in \mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \\ &\xlongequal{A_{ij} \in \mathcal{E}(\Omega)} -\sum_{i=1}^{N} \sum_{j=1}^{P(N,i)} m(A_{ij}) log_2(\frac{m(A_{ij})}{|\mathcal{E}(i)|}) \\ &\xlongequal{|\mathcal{E}(i)|=F(i)-1} -\sum_{i=1}^{N} \sum_{j=1}^{P(N,i)} m(A_{ij}) log_2(\frac{m(A_{ij})}{F(i)-1}) \\ &= H_r \end{aligned} \tag{35}$$

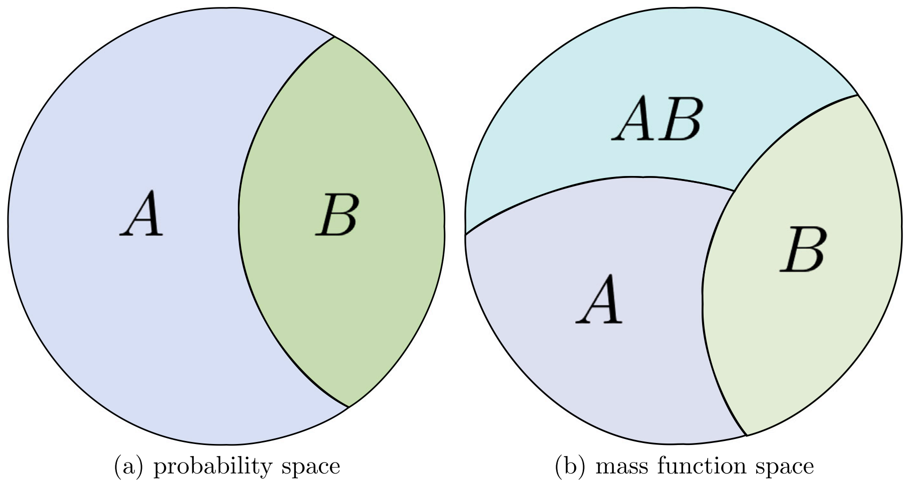


Fig. 2. Mass function assignment of sample space $\Omega = \{A, B\}$.

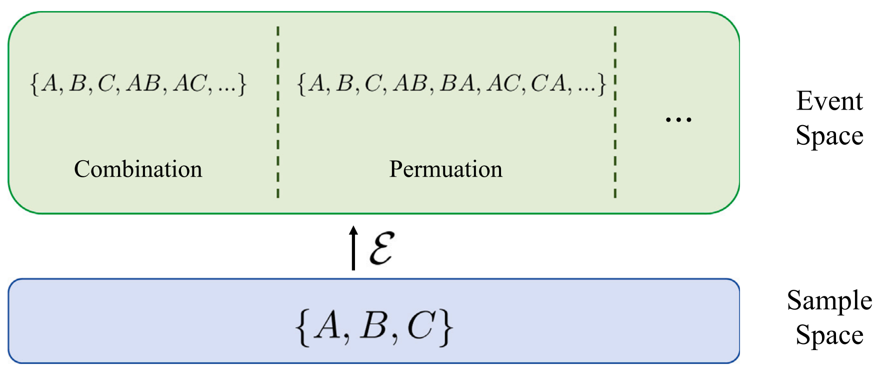


Fig. 3. Sample space $\Omega = \{A, B, C\}$ and different map $\mathcal{E}$ of mass function space $(\Omega, \mathcal{E}, m)$.

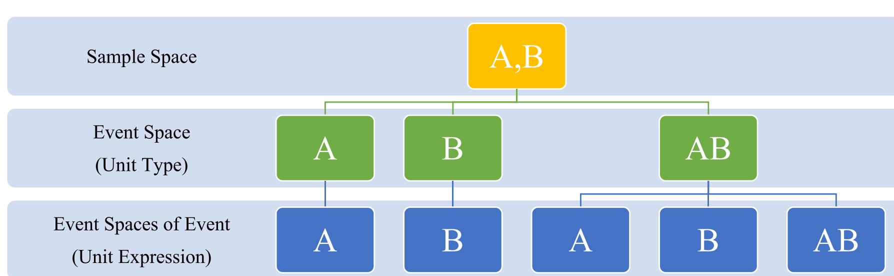


Fig. 4. Local similarity of GIE calculation and expression.

3.3. *Generalized information dimension*

Generalized Information Dimension (GID) D_g is defined as Eq. (36). For a certain infinite distribution of mass function, D_g is a constant. When mass function is probability mass function, GID degenerates to Rényi dimension [27]. GID is also fractal dimension of mass function distribution and reflect the increasing rate of information.

Definition 2 (*Generalized Information Dimension*).

$$D_g = \lim_{|\Omega| \to +\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \tag{36}$$

Example 6 (*Degeneration from GID to Rényi Dimension*). When mass function is probability and Ω is a discrete set, GID degenerates to Rényi Dimension in Eq. (37).

$$\begin{aligned} D_g &= \lim_{|\Omega| \to +\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \\ &\xlongequal[H_g = H_s]{|\mathcal{E}(\Omega)| = |\Omega|} \lim_{|\Omega| \to +\infty} \frac{H_s}{log_2 |\Omega|} \end{aligned} \tag{37}$$

The process function $\hat{D}_g$ is defined as Eq. (38). D_g in Eq. (36) is the numerical value at which the process function of $\hat{D}_g$ takes its limit. c is a constant in Eq. (38) which depends on event space $\mathcal{E}$ and distribution of mass function. c comes from removing the limitation operator in Eq. (36).

Definition 3 (*Process Function of Generalized Information Dimension*).

$$\hat{D}_g(n) = \hat{D}_g(|\Omega| = n) = \frac{H_g + c}{log_2(|\mathcal{E}(\Omega)|)} \tag{38}$$

3.4. An approximate value of entropy

Because GID D_g of a certain mass distribution is a constant, Approximate Generalized Information Entropy (AGIE) $\hat{H}_g$ is calculated in Eq. (39) which is a transformation of Eq. (38). As the size $|\Omega|$ of sample space increases, the computational complexity of GIE H_g can become enormous, and even encounter exponential explosions, making it impossible to obtain completely accurate data. Therefore, the accurate value of GID D_g can only be calculated through symbols inference rather than numerical values. Eq. (40) shows that $\hat{D}_g(\gamma)$ is consistent estimator of D_g. Then, in order to ensure the maximum accuracy, γ is set to the maximum sample space size that H_g can be calculated. The constant $c(\gamma)$ term of AGIE can be obtained in the Eq. (41). Some GIEs may have an exponential explosion in the size $|\mathcal{E}(\Omega)|$ of the event space, so AGIE can use reasonable estimator $|\mathcal{E}(\hat{\Omega})|$.

Definition 4 (*Approximate Generalized Information Entropy*).

$$\hat{H}_g(|\Omega|,\gamma) = \hat{D}_g(\gamma) * log_2(|\mathcal{E}(\hat{\Omega})|) + c(\gamma) \tag{39}$$

$$\lim_{|\gamma|\to+\infty} |\hat{D}_g(\gamma) - D_g| = 0 \tag{40}$$

$$\begin{aligned} c(\gamma) &= H_g - \hat{D}_g(\gamma) * log_2(|\mathcal{E}(\hat{\Omega})|) \\ &= -\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) - \hat{D}_g(\gamma) * log_2(|\mathcal{E}(\hat{\Omega})|) \end{aligned} \tag{41}$$

AGIE reveals that information entropy equals to increasing rate $\hat{D}_g$ multiply bit number of event space states $|\mathcal{E}(\Omega)|$. In the sight of Boltzmann entropy in Eq. (1), entropy is equal to constant k_b multiply natural bit number $ln(\cdot)$ of states $|W|$. Hence, the form and meaning of GIE in Eq. (39) is tend to Boltzmann Entropy in Eq. (1).

4. Properties

Next, some properties of GIE and GID will be introduced.

4.1. Maximum generalized information entropy

When sample space Ω and event space $\mathcal{E}(\Omega)$ is certain, the maximum GIE distribution is in Eq. (42).

$$m(i) = \frac{|\mathcal{E}(i)|}{\sum_{j\in\mathcal{E}(\Omega)} |\mathcal{E}(j)|} \tag{42}$$

Proof. GIE with Lagrange function [39] is in Eq. (43).

$$H_g = -\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) + \lambda(\sum_{i\in\mathcal{E}(\Omega)} m(i) - 1) \tag{43}$$

Gradient ∇H_g of Eq. (43) is Eq. (44).

$$\nabla H_{gi} = \frac{\partial H_g}{\partial m(i)} = -log_2 \frac{m(i)}{|\mathcal{E}(i)|} - \frac{1}{ln2} + \lambda \tag{44}$$

Make the gradient equal to 0 to solve for the maximum value, while combining the constraints of the mass function in Eq. (45).

$$\begin{cases} \nabla H_g = 0 \\ \sum_{i\in\mathcal{E}(\Omega)} m(i) = 1 \end{cases} \tag{45}$$

Solve of Eq. (45) is maximum GIE distribution in Eq. (42). The value of maximum GIE is in Eq. (46).

$$Max(H_g) = log_2(\sum_{i\in\mathcal{E}(\Omega)} |\mathcal{E}(i)|) \quad \square \tag{46}$$

Example 7 (*Maximum GIE of Rare Event Space*). Assume the sample space is defined in Eq. (47). And the event space $\mathcal{E}(\cdot)$ is generated by followed rules:

1. If $|i| = 1$, $\mathcal{E}(i) = i$

Table 1
Example event space pattern.

Event i	Event space $\mathcal{E}(i)$	Size $\lvert\mathcal{E}(i)\rvert$	$m(i)$
A	$\{A\}$	1	$\frac{1}{9}$
B	$\{B\}$	1	$\frac{1}{9}$
C	$\{C\}$	1	$\frac{1}{9}$
AB	$\{\emptyset, A, B, \cancel{AB}\}$	2	$\frac{2}{9}$
AC	$\{\emptyset, A, C, \cancel{AC}\}$	2	$\frac{2}{9}$
BC	$\{\emptyset, B, C, \cancel{BC}\}$	2	$\frac{2}{9}$
$\cancel{ABC}$	$\{\emptyset, A, B, C, AB, AC, BC, \cancel{ABC}\}$	6	/

2. If $|i| > 1$, $\mathcal{E}(i) = 2^i - i - \emptyset$

And for sample space Ω, the corresponding event spaces are shown in Table 1.

$$\Omega = \{A, B, C\} \tag{47}$$

The event space $\mathcal{E}$ is able to be regarded as remove the sample space from its power set and when sample space only contains one sample, event space is sample space itself. Then the GIE is as Eq. (48).

$$\begin{aligned} H_g &= -\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \\ &= -m(A) * log_2(m(A)) + m(B) * log_2(m(B)) + \\ &m(C) * log_2(m(C)) + m(AB) * log_2(\frac{m(AB)}{2}) + \\ &m(AC) * log_2(\frac{m(AC)}{2}) + m(BC) * log_2(\frac{m(BC)}{2}) \end{aligned} \tag{48}$$

When GIE comes the maximum value, the maximum condition is shown in Table 1 which is consistent with Eq. (42). The value of the maximum GIE is in Eq. (49) which is consistent with Eq. (46).

$$\begin{aligned} H_g &= -(3 * \frac{1}{9} * log_2(\frac{1}{9}) + 3 * \frac{2}{9} * log_2(\frac{2}{9 * 2})) \\ &= log_2 9 \end{aligned} \tag{49}$$

4.2. Finite distribution

Given a finite mass distribution, corresponding GID is 0 in Eq. (50) and corresponding GIE is a constant in Eq. (51). Rényi dimension and Shannon entropy have the same property [27].

$$D_g = 0 \tag{50}$$

$$H_g = c \tag{51}$$

4.3. Infinite distribution: Process view

Mass function is able to follow arbitrary infinite distribution. Thus, with the same distribution, there is a process of mass functions with increasing $|\mathcal{E}(\Omega)|$.

There are several common distributions in Eqs. (52), (53), (54), (55) and (56). The significance of mentioned distributions are listed in Table 2.

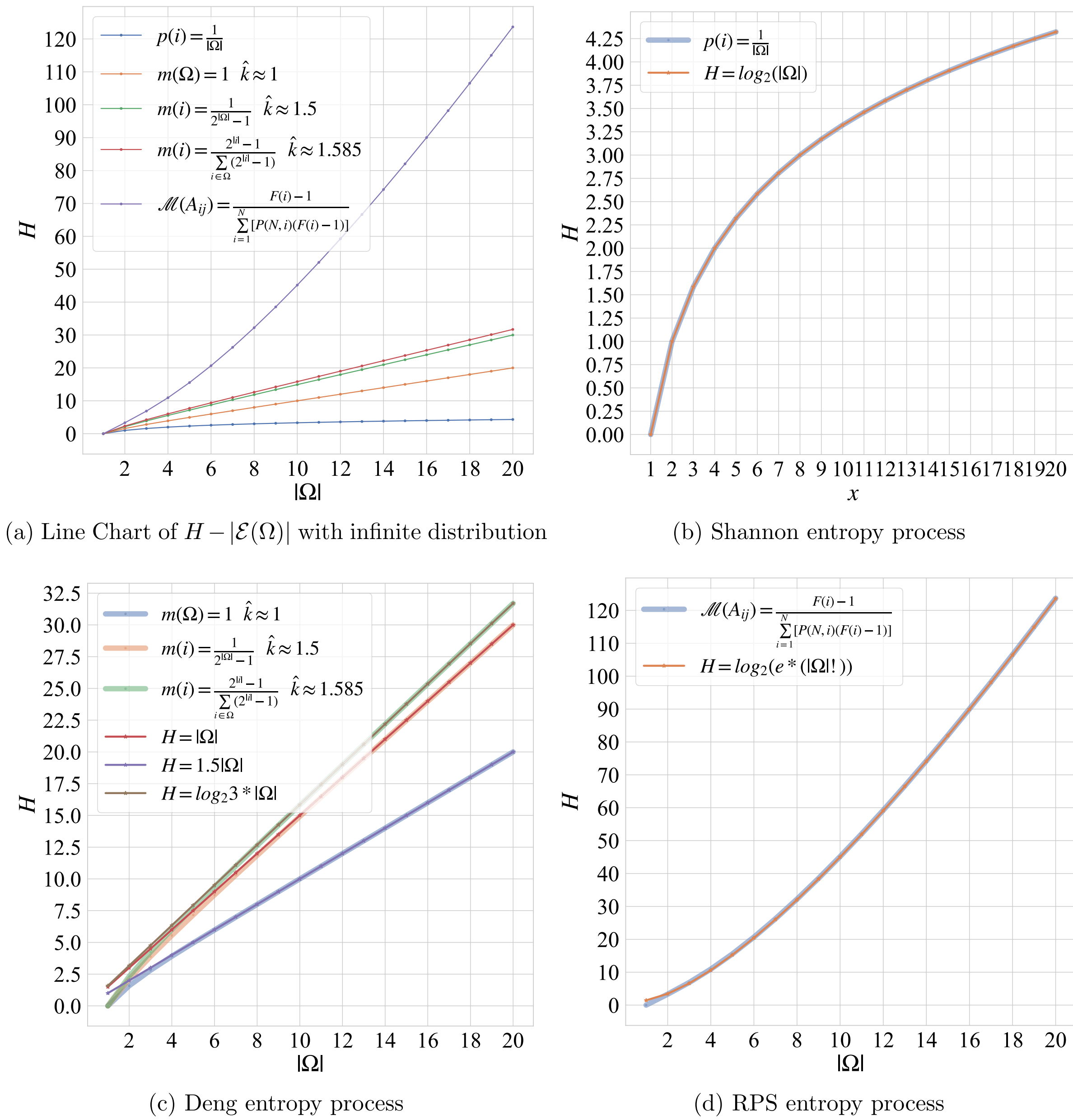
$$p(i) = \frac{1}{|\Omega|} \tag{52}$$

$$m(\Omega) = 1 \tag{53}$$

$$m(i) = \frac{1}{2^{|\Omega|} - 1} \tag{54}$$

$$m(i) = \frac{2^{|i|} - 1}{\sum_{i\in\Omega} (2^{|i|} - 1)} \tag{55}$$

$$\mathscr{M}(A_{ij}) = \frac{F(i) - 1}{\sum_{i=1}^{N} [P(N,i)(F(i) - 1)]} \tag{56}$$



(a) Line Chart of $H - |\mathcal{E}(\Omega)|$ with infinite distribution

(b) Shannon entropy process

(c) Deng entropy process

(d) RPS entropy process

Fig. 5. GIE process with increasing $|\mathcal{E}(\Omega)|$.

Table 2
The significance of selected distribution.

Equation number	Significance
(52)	Maximum Shannon entropy
(53)	Full assignment to single event
(54)	Uniform assignment
(55)	Maximum Deng entropy
(56)	Maximum RPS entropy

The process lines of GIE from $|\Omega| = 1$ to $|\Omega| = 20$ are plotted in Fig. 5(a). To clearly analyze the law of variation, lines of Shannon entropy, Deng entropy and RPS entropy are split into Figs. 5(b)–5(d).

4.3.1. Shannon entropy process

The process line of Shannon entropy is shown in Fig. 5(b). As the expanding of sample space, the process line exhibits a logarithmic function characteristic. As discussed in the relationship between mass and probability, the events in the event space of Shannon entropy are composed of a single sample. In other words, the event space of Shannon entropy is the same as the sample space Ω in Eq. (57). When the probability is distributed from the maximum entropy, the formula of GID D_g corresponding to the Shannon entropy is a constant 1 in Eq. (58). According to the calculation method of AGIE, the corresponding constant terms can be calculated in Eq. (59). Because both the GID and constant terms here are accurate values, the analytical expression of AGIE $\hat{H}_g$ is independent of the sample space size γ for maximum accuracy. At this point, there is no error in AGIE, which equals GIE in Eq. (60).

$$|\mathcal{E}(\Omega)| = |\Omega| \tag{57}$$

$$\begin{aligned} D_g &= \lim_{|\Omega|\to+\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{log_2(|\Omega|)}{log_2(|(\Omega)|)} \\ &= 1 \end{aligned} \tag{58}$$

$$c = log_2(|\Omega|) - log_2(|\Omega|) = 0 \tag{59}$$

Table 3
Estimated slope and GID of Deng entropy process line.

Equation number	Estimated slope	GID
(53)	1	1
(54)	1.5	1.5
(55)	1.585	$log_2 3 \approx 1.585$

The curves of AGIE and GIE completely overlap in Fig. 5(b). Because the size of the Shannon entropy event space is consistent with its sample space, it ultimately presents in logarithmic form. When the probability distribution is not the maximum entropy, the GID and constant terms varies. When GID is not equal to zero, Shannon entropy increases logarithmically with the increasing size of sample space, and only the coefficients $\hat{D}_g(\gamma)$ of logarithmic and constant terms $c(\gamma)$ change in Eq. (61).

$$\begin{aligned} &\hat{H}_g = log_2(|\mathcal{E}(|\Omega|)) * \hat{D}_g + c \\ \xRightarrow[|\mathcal{E}(\Omega)|=|\Omega|]{D_g=1} &\hat{H}_g = log_2(|\Omega|) = H_g \end{aligned} \tag{60}$$

$$\hat{H}_g(|\Omega|, \gamma) = \hat{D}_g(\gamma) * log_2(|\Omega|) + c(\gamma) \tag{61}$$

4.3.2. Deng entropy process

Zhao et al. proposed the linearity of Deng entropy [40]. The fundamental law of linearity phenomenon is able to explained by GIE and GID. In Fig. 5(c), there is truly a phenomenon that Deng entropy increases linearly with the size of the sample space. The event space $\mathcal{E}$ of Deng entropy is composed of a combination of sample spaces Ω, so the size $|\mathcal{E}(\Omega)|$ of event space is in Eq. (62). To avoid exponential explosion, the estimated size $|\mathcal{E}(\hat{\Omega})|$ of the event space is shown in Eq. (63). $|\mathcal{E}(\hat{\Omega})|$ is not a consistent estimator, but $log_2(|\mathcal{E}(\hat{\Omega})|)$ is a consistent estimator in Eq. (64). So the approximate entropy of Deng entropy can be written as Eq. (65). Therefore, the Deng entropy process exhibits linearity.

$$|\mathcal{E}(\Omega)| = 2^{|\Omega|} - 1 \tag{62}$$

$$|\mathcal{E}(\hat{\Omega})| = 2^{|\Omega|} \tag{63}$$

$$\lim_{|\Omega|\to+\infty} log_2(2^{|\Omega|} - 1) - log_2(2^{|\Omega|}) = \lim_{|\Omega|\to+\infty} log_2(\frac{2^{|\Omega|} - 1}{2^{|\Omega|}}) = 0 \tag{64}$$

$$\begin{aligned} \hat{H}_g(|\Omega|, \gamma) &= \hat{D}_g(\gamma) * log_2(2^{|\Omega|}) + c(\gamma) \\ &= \hat{D}_g(\gamma) * |\Omega| + c(\gamma) \end{aligned} \tag{65}$$

The first three curves in Fig. 5(c) represent the actual values of Deng entropy under different distributions, while the $\hat{k}$ in the legend represents the slope of the actual curve obtained by the least squares method. Referring to Eq. (65) the estimate of slope $\hat{k}$ is approximate to GID. The GID of Deng entropy can be accurately calculated. Hence, GIDs of selected three mass distributions are show in Table 3. Estimated slopes of Eqs. (53) and (54) is equal to corresponding GIDs, and due to $|\Omega|$ is not large enough, the approximate value of estimated slope $\hat{k} = 1.585$ is approximate to GID $D_g = log_2 3$ in Eq. (55).

The last three curves in Fig. 5(c) are AGIE of selected distributions. The curves of AGIE and GIE do not completely overlap in Fig. 5(c). When $|\Omega| \in [1, 3]$, there is a certain deviation between AGIE and Deng entropy. This deviation originates from the feature of approximation, where the error between the estimated value and the target value decreases as increasing of $|\Omega|$, and eventually converges to 0. So, when $|\Omega|$ is a small number, it is not recommended to calculate by AGIE, but to directly calculate the exact value of GIE.

4.3.3. RPS entropy process

There is an exponential explosion problem with the size $|\mathcal{E}(\Omega)|$ of the event space in RPS that $|\mathcal{E}(\Omega)|$ cannot accurately calculate timely when size $|\Omega|$ of sample space is a big number. To solve this problem, Zhou et al. proposed the maximum envelope $\lim_{|\Omega|\to+\infty} \sum_{i\in\mathcal{E}(\Omega)} |\mathcal{E}(i)|$ of RPS under the limit condition in Eq. (66) [41]. So according to the maximum entropy form of GIE in Eq. (46), approximate RPS entropy can be calculated as Eq. (67).

$$\lim_{|\Omega|\to+\infty} \sum_{i\in\mathcal{E}(\Omega)} |\mathcal{E}(i)| = \lim_{|\Omega|\to+\infty} e * (|\Omega|!)^2 \tag{66}$$

$$\hat{H}_g(|\Omega|) = log_2(e * (|\Omega|!)^2) \tag{67}$$

In Fig. 5(d), the two curves represent the accurate value of GIE and the value of AGIE. Similar to the process of Dentropy, when $|\Omega| \in [1, 3]$, there is a certain deviation between AGIE and GIE. This phenomenon comes from the fact that AGIE converges to GIE under extreme conditions. When the sample space is too small, the gap will be obvious, so a hybrid calculation strategy of GIE is proposed to optimize the problem in Section 5.3.

4.4. GIE of random distributions

There are infinity cases of distribution. The distributions mentioned in Section 4.3 are common. Special two-point distribution $\kappa(m, n)$ is shown as an example of random distributions which is defined as follows:

- The sample space size $|\Omega|$ is equal to n. Randomly choose two events E_1, E_2 which size is equal to $|\Omega| - 1$ and assign $(m, 1 - m)$ as their mass functions. The event space is defined as power set. Hence, the GID of $\kappa(m, 1 - m)$ is in Eq. (68).

$$\begin{aligned} D_g &= \lim_{|\Omega|\to+\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-m * log_2(\frac{m}{2^{|\Omega|-1}-1}) - (1 - m) * log_2(\frac{1-m}{2^{|\Omega|-1}-1})}{log_2(2^{|\Omega|} - 1)} \\ &= \lim_{|\Omega|\to+\infty} \frac{log_2(2^{|\Omega|-1} - 1)}{log_2(2^{|\Omega|} - 1)} \\ &= \lim_{|\Omega|\to+\infty} \frac{2^{|\Omega|} - 1}{2^{|\Omega|} - 2} \\ &= 1 \end{aligned} \tag{68}$$

In the case, m and GID is not relevant. According to Eq. (39), the approximate curves and process line when $m = \frac{1}{5}$ are plotted in Fig. 6. There is a constant bias $c = 0.278$ of approximate curve. And for saving performance of computer, there is an approximate algorithm of GIE process value in Section 5.3.

5. Applications

5.1. Binary search

After numerical example, there is a practical application in Application 1 to verify proposed entropy and dimension.

Application 1 (*Single No. 1 Problem*)**.** *In a game with 32 teams, the organizer has access to all their orders. In order to identify the single top 1 team (such a competition form) who have the highest order, how many questions must we ask at most? The organizer can only respond with "Yes" or "No".*

Solve of Application 1:

1. Identify sample space Ω as Eq. (69). A_i stands for "ith participant gets the highest score".

$$\Omega = \{A_1, A_2, A_3, \dots, A_{32}\} \tag{69}$$

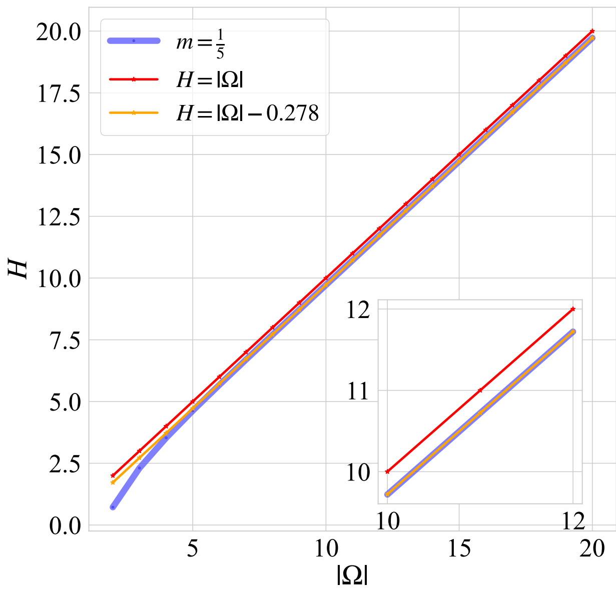


Fig. 6. Special two-point distribution when $m = \frac{1}{5}$.

2. Identify event space $\mathcal{E}(\Omega)$ as Eq. (70).

$$\mathcal{E}(\Omega) = \Omega \tag{70}$$

3. Assign the general mass function as Eq. (71) because there are 1024 different units in state space $\mathcal{E}(\Omega)$.

$$m(i) = \frac{1}{|\mathcal{E}(\Omega)|} = \frac{1}{32},\ i \in \mathcal{E}(\Omega) \tag{71}$$

4. Calculate the Generalized Information Dimension according to Distribution as Eq. (72).

$$\begin{aligned} D_g &= \lim_{|\Omega|\to+\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|})}{log_2(|\Omega|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-\sum^{|\Omega|} \frac{1}{|\Omega|} * log_2(\frac{\frac{1}{|\Omega|}}{1})}{log_2(|\Omega|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{log_2(|\Omega|)}{log_2(|\Omega|)} \\ &= 1 \end{aligned} \tag{72}$$

5. Calculate the Number of problems N which is equal to Generalized Information Entropy as Eq. (73).

$$\begin{aligned} N &= \lceil D_g * log_2(|\mathcal{E}(\Omega)|) \rceil \\ &= \lceil 1 * log_2(32) \rceil \\ &= 5 \end{aligned} \tag{73}$$

5.2. Ergodicity

Application 2 (*Group of Highest Score Participant Problem*)**.** *In a test with 32 participants, the organizer has access to all their scores. In order to identify the group of top 1 participant (such as enrollment exam) who have the highest score, how many questions must we ask at most? The organizer can only respond with "Yes" or "No".*

Solve of Application 2:

1. Identify sample elements x as Eq. (69) which is same as Application 1. A_i stands for "ith participant gets the highest score".
2. Identify event space $\mathcal{E}(\Omega)$ as Eq. (74). $\mathcal{E}(\Omega)$ is the power set of x because how many participants get the highest score is unknown.

$$\mathcal{E}(\Omega) = \{\{A_1\},\{A_2\},\ldots,\{A_{32}\},\{A_1,A_2\},\ldots,\Omega\} \tag{74}$$

3. Assign the general mass function as Eq. (75) because all participants are possible to get the highest score.

$$m(\Omega) = m(A_1, A_2, \ldots, A_{32}) = 1 \tag{75}$$

4. Calculate the Generalized Information Dimension according to Distribution as Eq. (76).

$$\begin{aligned} D_g &= \lim_{|\Omega|\to+\infty} \frac{H_g}{log_2(|\mathcal{E}(\Omega)|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|})}{log_2(|\mathcal{E}(\Omega)|)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-m(x) * log_2(\frac{m(x)}{2^{|\Omega|}-1})}{log_2(2^{|\Omega|}-1)} \\ &= \lim_{|\Omega|\to+\infty} \frac{-1 * log_2(\frac{1}{2^{|\Omega|}-1})}{log_2(2^{|\Omega|}-1)} \\ &= \lim_{|\Omega|\to+\infty} \frac{log_2(2^{|\Omega|}-1)}{log_2(2^{|\Omega|}-1)} \\ &= 1 \end{aligned} \tag{76}$$

5. Calculate the Number of problems N which is equal to Generalized Information Entropy as Eq. (77).

$$\begin{aligned} N &= \lceil D_g * log_2(|\mathcal{E}(\Omega)|) \rceil \\ &= \lceil 1 * log_2(2^{32}-1) \rceil \\ &= 32 \end{aligned} \tag{77}$$

Application 1 is classic problem of binary search which contains single target. And Application 2 is classic problem of ergodicity which contains unknown number of targets.

5.3. Estimation of entropy under large number

In the sight of GIE, there is a heavy problem that if the size of the sample space is large, GIE may not be computable. In the other hand, single computer performance is not enough to calculate the accurate value of GIE, and approximate value is able to provide a relative accurate value of GIE. In Section 3.4, the specific form of AGIE is proposed. There may be a certain gap between the approximate value and GIE, but it will gradually converge to 0 as size $|\Omega|$ of sample space increases. This phenomenon is embodied in the Deng entropy process and the RPS process in Sections 4.3.2 and 4.3.3. Then the accuracy of GIE can be guaranteed to the greatest extent through the hybrid strategy:

1. When size $|\Omega|$ is computable, to ensure the most accuracy, GIE in Eq. (26) is the best way to calculate GIE.
2. When size $|\Omega|$ is not compute, to ensure the most accuracy, AGIE in Eq. (39) is the best way to calculate GIE.

In the specific calculation of AGIE, the determining size γ of the maximum computable sample space needs to be completed first in Algorithm 1. So, the generated GIE $H(i), i \in [1,\gamma]$ can be used for least squares estimation of GID $\hat{D}_g(\gamma)$ and constant $c(\gamma)$.

Another approximate scenario for calculating GIE is to specify precision σ. When precision is specified, estimating GID and constant terms does not require reaching the maximum computable sample space. When the approximate GID converges within the specified precision, AGIE will meet the conditions for the specified precision.

According to Eq. (39), approximate GIE $\hat{H}_g$ calculation needs approximate GID $\hat{D}_g$ and bias c. Algorithm 2 is a swift way to make

Algorithm 1: Determine the largest size of computable sample space

Input: Generalized Information Entropy $H_g(\gamma)$ with sample space size γ under a specified distribution
Output: Computable maximum sample space size γ

```
// Initialize sample space size
γ = 1;
while True do
    // Determine whether GIE can be calculated at this time
    if H_g(γ) then
        // Update maximum sample space size
        γ = γ + 1;
    else
        break;
    end
end
return γ
```

sure the approximate GID $\hat{D}_g$ and bias c at current accuracy σ. So the corresponding AGIE calculation algorithm is in Algorithm 2.

Algorithm 2: Estimate Generalized Information Dimension and constant item

Input: Generalized information entropy $H_g(n)$ with sample space size n under a specified distribution, target precision σ, function of event space size $|\mathcal{E}|(n)$
Output: Approximate generalized information dimension $\hat{D}_g$, constant item of approximate generalized information entropy c

```
// Initialize the primary size of sample space
n = 2;
// Initialize the primary Generalized Information Dimension
D_pre = H(n) / log_2(|E|(n));
// Traverse size of sample space
while True do
    // Update the size of sample space
    n = n + 1;
    // Update the Generalized Information Dimension
    D_cur = H(n) / log_2(|E|(n));
    // Determine whether the target accuracy has been achieved
    if |D_cur − D_pre| ⩽ σ then
        D̂_g = D_cur;
        break;
    end
end
// Calculate the bias of approximate Generalized Information Entropy
c = H(n) − D̂_g ∗ log_2(|E|(n));
return D̂_g, c
```

Example in Section 4.4 is generated by Algorithm 2. The last two curves in Fig. 6 are both AGIE curves, and the difference between the two curves is the presence or absence of a constant term c. Accurate GID $\hat{D}_g$ is in Eq. (68) and after estimation of $c = -0.2780$, approximate GIE is in Eq. (78). In order to analyze the gap of AGIE and GIE more clearly, the difference between GIE and AGIE is shown in Fig. 7. It can be observed that as the sample space increases, GIE and AGIE completely overlap, with an error of 0, and the estimation effect is ideal.

$$\begin{aligned} \hat{H}_g &= \hat{D}_g * log_2(|\mathcal{E}(x)|) + c \\ &\xlongequal[c=-0.278]{D_g=1} log_2(2^{|\Omega|} - 1) - 0.278 \\ &\approx |\Omega| - 0.278 \end{aligned} \tag{78}$$

5.4. Information representation

Information entropy represents the average amount of information. Shannon entropy can be written as the expectation of the function $I(x)$ in Eqs. (79) and (80) where $I(x)$ is information of event x and $\mathcal{E}$ is event space.

$$I(x) = -log_2(p(x)) \tag{79}$$

$$H_s = E_p(I(x)) = -\sum_{x\in\mathcal{E}} p(x) * log_2(p(x)) \tag{80}$$

The unit of information storage corresponding to Shannon entropy is bit, and there is no difference between bits. When the bit expresses 0, it indicates that nothing has happened, and 1 indicates that it has occurred, which is the same as the meaning of Boolean algebra. To distinguish different significance of bit, a fixed-length encoding is common to used. A fixed-length code consists of n bits, and represent 2^n states. The feature of fixed-length encoding is also obvious that different events can only be expressed by combining bits. The left side of Fig. 8 is a fixed-length encoding for the 5 objects.

From the perspective of mass function, the information storage units corresponding to each event in the event space are different. The states that each unit can express are consistent with the event space generated by the corresponding event. Therefore, in GIE, information of event x is defined as Eq. (81) and GIE can also be written in form of expectation in Eq. (82). Fig. 4 is an example of encoding expression of sample space $\{A, B\}$. 3 different storage units were generated, expressing up to 5 different states.

$$I(x) = -log_2(\frac{m(x)}{|\mathcal{E}(x)|}) \tag{81}$$

$$H_g = E_m(I(x)) = -\sum_{i\in\mathcal{E}(\Omega)} m(i) * log_2(\frac{m(i)}{|\mathcal{E}(i)|}) \tag{82}$$

In GIE coding, distinguishing a object requires distinguishing both the information unit and the state of the information unit. It is also consistent with the physical meaning of mass function. Different information units are analogized to different particles. The encoding of GIE does not require the combination of bits. In other words, each state corresponds to one state of one information unit. The right side of Fig. 8 is the case of Deng entropy encoding. Different colored boxes represent different information units and events at the same time. And the numbers in the box represent the samples in the sample space. The blue box represents event 0, the yellow box represents event 1, and the green box represents event 01. It is worth noting that the box corresponding to event 01 can only express 0 or 1, and in Fig. 8, it corresponds to Object 3 and Object 4. The meaning of maximum entropy is that all states expressed by information units are effectively utilized, so the frequency of each state is equal. This is also the same as the condition for maximum GIE in Eq. (42) that the mass function of an event is proportional to the number of states that the information unit of the event can express.

6. Conclusion

By observing and analyzing the formal changes from Boltzmann entropy to Shannon entropy and Deng entropy, this paper proposes a generalized information entropy to accommodate different forms of information measurement and discover the common properties of different forms of information entropy. Information dimension is a fractal dimension that describes the growth rate of information. To

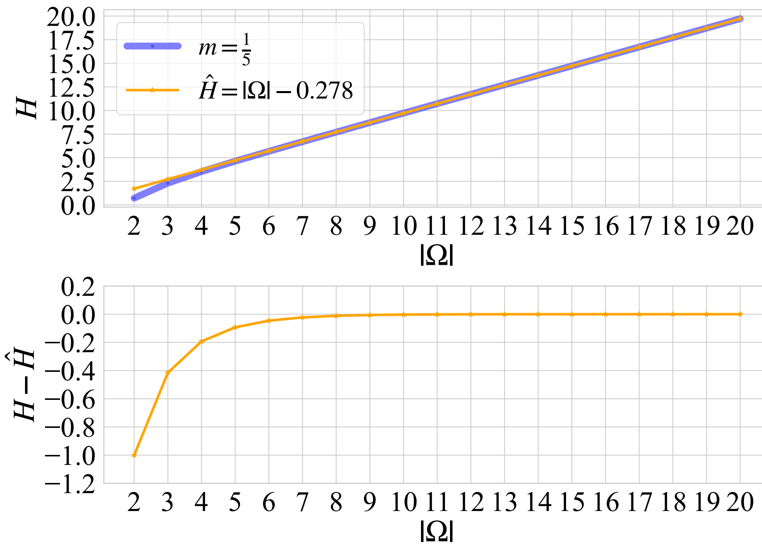


Fig. 7. Loss between approximate GIE and accurate GIE.

Fix-length Encoding

Objects	Code
Object 1	0 0 0
Object 2	0 0 1
Object 3	0 1 0
Object 4	0 1 1
Object 5	1 0 0

GIE Encoding

Objects	Code
Object 1	0
Object 2	1
Object 3	0
Object 4	1
Object 5	01

Fig. 8. Different encoding patterns between Shannon entropy and Deng entropy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

accommodate different forms of information dimension, this paper also proposes a generalized information dimension. The most important conclusion is that information entropy can be determined through information dimension and event space.

In practical applications, generalized information entropy has made important contributions to the approximate calculation of entropy and the encoding logic corresponding. Through the relationship between generalized information entropy and generalized information dimension, generalized information entropy can be calculated in an approximate form under the condition of large numbers, effectively avoiding the problem of incalculability caused by exponential explosion. Meanwhile, a hybrid computing strategy has also been proposed to ensure maximum accuracy. Another application of generalized information entropy is encoding systems. In the perspective of generalized information entropy, the representation of information has also been expanded, not limited to bits. Mass function makes events particle-like, and the same event can have different representation states. Therefore, the generalized information entropy exhibits a similar form to Boltzmann entropy at the limit state.

The other properties of generalized information entropy and generalized information dimensionality are still being explored. The information representation in generalized information entropy shares similarities with quantum theory, and the potential correlations need to be further studied. On the other hand, generalized information entropy considers discrete systems and is not effectively compatible with continuous information measurement. Further research is being conducted on the representation and measurement methods of continuous information. Generalized information entropy is a type of entropy with more physical properties and contains potential.

CRediT authorship contribution statement

Tianxiang Zhan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Jiefeng Zhou:** Resources, Methodology, Conceptualization. **Zhen Li:** Writing – review & editing, Supervision, Project administration. **Yong Deng:** Writing – review & editing, Investigation, Funding acquisition, Project administration, Resources, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The work is partially supported by National Natural Science Foundation of China (Grant No. 62373078). The authors thank Qianli Zhou (Ph.D. candidate from University of Electronic Science and Technology of China) for discussing significance of event space. The authors also would like to appreciate Zihan Yu (master candidate from University of Electronic Science and Technology of China) for his contributions in discussing the similarities and differences between Shannon entropy and Deng entropy.

References

[1] Gauthier J, Derome N. Evenness-richness scatter plots: a visual and insightful representation of shannon entropy measurements for ecological community analysis. Msphere 2021;6(2):10–128.
[2] Li W, Xie D, Li D, Zhang Y, Gao Y, Liaw PK. Mechanical behavior of high-entropy alloys. Prog Mater Sci 2021;118:100777.
[3] Mao A, Mohri M, Zhong Y. Cross-entropy loss functions: Theoretical analysis and applications. In: International conference on machine learning. PMLR; 2023, p. 23803–28.
[4] Clausius R. Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie: vorgetragen in der naturforsch. Gesellschaft den 24. April 1865. éditeur inconnu; 1865.
[5] Perrot P. A to Z of thermodynamics. Supplementary Series; 27; 1998.
[6] Jaynes ET. Gibbs vs Boltzmann entropies. Am J Phys 1965;33(5):391–8.
[7] Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27(3):379–423.
[8] Kharazmi O, Contreras-Reyes JE, Balakrishnan N. Jensen–Fisher information and Jensen–Shannon entropy measures based on complementary discrete distributions with an application to Conway's game of life. Physica D 2023;133822.
[9] Lin J. Divergence measures based on the Shannon entropy. IEEE Trans Inf Theory 1991;37(1):145–51.
[10] Flamich G, Markou S, Hernández-Lobato JM. Fast relative entropy coding with a* coding. In: International conference on machine learning. PMLR; 2022, p. 6548–77.
[11] Flamich G, Havasi M, Hernández-Lobato JM. Compressing images by encoding their latent representations with relative entropy coding. Adv Neural Inf Process Syst 2020;33:16131–41.
[12] Ricci L, Perinelli A, Castelluzzo M. Estimating the variance of Shannon entropy. Phys Rev E 2021;104(2):024220.
[13] Cincotta PM, Giordano CM. Estimation of diffusion time with the Shannon entropy approach. Phys Rev E 2023;107(6):064101.
[14] Cover TM. Elements of information theory. John Wiley & Sons; 1999.
[15] Jaynes ET. Information theory and statistical mechanics. Phys Rev 1957;106(4):620.
[16] Jaynes ET. Information theory and statistical mechanics. II. Phys Rev 1957;108(2):171.
[17] Jaynes ET. Prior probabilities. IEEE Trans Syst Sci Cybern 1968;4(3):227–41.
[18] Zivieri R. Magnetic skyrmions as information entropy carriers. IEEE Trans Magn 2022;58(1):1–5.
[19] Zivieri R. From thermodynamics to information: Landauer's limit and negentropy principle applied to magnetic skyrmions. Front Phys 2022;10:769904.
[20] Dempster A. Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 1967;38:325–39.
[21] Shafer G. A mathematical theory of evidence. Vol. 42, Princeton University Press; 1976.
[22] Deng Y. Deng entropy. Chaos Solitons Fractals 2016;91:549–53.
[23] Liao H, Ren Z, Fang R. A deng-entropy-based evidential reasoning approach for multi-expert multi-criterion decision-making with uncertainty. Int J Comput Intell Syst 2020;13(1):1281–94.
[24] Kharazmi O, Contreras-Reyes JE. Deng–Fisher information measure and its extensions: Application to Conway's game of life. Chaos Solitons Fractals 2023;174:113871.
[25] Contreras-Reyes JE, Kharazmi O. Belief Fisher–Shannon information plane: Properties, extensions, and applications to time series analysis. Chaos Solitons Fractals 2023;177:114271.
[26] Deng Y. Random permutation set. Int J Comput Commun Control 2022;17(1).
[27] Rényi A. On the dimension and entropy of probability distributions. Acta Math Acad Sci Hung 1959;10(1–2):193–215.
[28] Geiger BC, Koch T. On the information dimension of stochastic processes. IEEE Trans Inf Theory 2019;65(10):6496–518.
[29] Wu Y, Verdú S. Rényi information dimension: Fundamental limits of almost lossless analog compression. IEEE Trans Inform Theory 2010;56(8):3721–48.
[30] Wu Y, Shitz SS, Verdú S. Information dimension and the degrees of freedom of the interference channel. IEEE Trans Inform Theory 2014;61(1):256–79.
[31] Qiang C, Deng Y, Cheong KH. Information fractal dimension of mass function. Fractals 2022;30(06):2250110.
[32] Zhao T, Li Z, Deng Y. Information fractal dimension of random permutation set. Chaos Solitons Fractals 2023;174:113883.
[33] Halpern JY, Fagin R. Two views of belief: belief as generalized probability and belief as evidence. Artif Intell 1992;54(3):275–317.
[34] Ryser HJ. Combinatorial mathematics. Vol. 14, American Mathematical Soc.; 1963.
[35] Vicsek T. Fractal growth phenomena. World scientific; 1992.
[36] Sagan H. Space-filling curves. Universitext 1994.
[37] Falconer K. Fractal geometry: mathematical foundations and applications. John Wiley & Sons; 2004.
[38] Nguyen HT. On random sets and belief functions. J Math Anal Appl 1978;65(3):531–42.
[39] Beavis B, Dobbs I. Optimisation and stability theory for economic analysis. Cambridge University Press; 1990.
[40] Zhao T, Li Z, Deng Y. Linearity in Deng entropy. Chaos Solitons Fractals 2024;178:114388.
[41] Zhou J, Li Z, Cheong KH, Deng Y. Limit of the maximum random permutation set entropy. 2024, arXiv preprint arXiv:2403.06206.