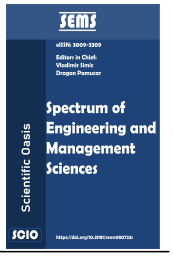
SCIENTIFIC OASIS

Spectrum of Engineering and Management Sciences

Journal homepage: www.sems-journal.org
ISSN: 3009-3309



Entropy Reimagined: Theoretical Foundations of HyperEntropy and SuperHyperEntropy with Real-World and Engineering Applications

Takaaki Fujita[1,*]

[1] Independent Researcher, Shinjuku, Shinjuku-ku, Tokyo, Japan.

ARTICLE INFO

Article history:
Received 13 April 2025
Received in revised form19 May 2025
Accepted 14 July 2025
Available online 19 July 2025

Keywords:
Entropy, HyperEntropy, SuperHyperEntropy, Hyperstructure, Superhyperstructure

ABSTRACT

Classical mathematical structures can be systematically extended into hyperstructures and superhyperstructures via the powerset and its n-fold iteration. These enriched frameworks are particularly well suited for modeling hierarchical or multi-layered relationships across diverse domains. Meanwhile, entropy quantifies the average information-theoretic uncertainty of a discrete probability distribution. In this paper, we introduce two new constructs: *HyperEntropy*, which extends Shannon entropy to distributions over subsets of a base set, and *SuperHyperEntropy*, which further generalizes the concept to distributions on iterated powerset levels. We develop their fundamental properties, present illustrative examples, and outline potential applications, thereby laying the groundwork for future advances in entropy theory. Additionally, we present a preliminary example of their application in engineering.

1. Introduction

1.1 Entropy

Entropy measures the average uncertainty, in the information-theoretic sense, of a discrete random variable's outcome distribution. As a fundamental concept, it underpins methods and analyses in areas such as machine learning, probability theory, computer science, physics, information theory, and optimization (see, e.g., [(1)–(4)]).

*Corresponding author.
E-mail address: Takaaki.fujita060@gmail.com

https://doi.org/10.31181/sems31202555t

1.2 Superhyperstructure

A *classical structure* denotes any mathematical framework—such as graphs, topologies, functions, algebras, groups, or automata—in which operations act directly on elements of a base set. A *hyperstructure* extends this notion by allowing its operations to take subsets of the base set as inputs and return subsets as outputs (cf.[(5)–(7)]). Pushing this idea further, an n-*superhyperstructure* applies the powerset construction iteratively n times, thereby modeling multi-level or hierarchical relationships among elements [(8)–(11)]. Well-known instances of such systems include *superhypergraphs* [(12)–(16)]. Both hyperstructures and their n-fold extensions furnish powerful abstractions for capturing layered or composite phenomena in complex systems (see Table 1).

Table 1
Overview of classical structures, hyperstructures, and n-superhyperstructures

Structure	Underlying Set	Operation	Key Feature
Classical Structure	S	$\#\colon S^m \to S$	A traditional algebraic system in which each operation on m elements yields a single element of S. Examples include groups, rings, and fields.
Hyperstructure	$\mathcal{P}(S)$	$\circ\colon S \times S \to \mathcal{P}(S)$	A set-valued generalization: combining two elements produces a nonempty subset of S, enabling multivalued or collective behaviors.
n-Superhyperstructure	$P_n(S)$	$\circ^{(m,n)}\colon S^m \to \mathcal{P}^n(S)$	A hierarchical extension using the n-fold powerset: operations map m elements into the n-th iterated powerset, modeling nested, multilayered relationships.

1.3 Our Contribution: HyperEntropy and SuperHyperEntropy

Although entropy theory and (super)hyperstructural frameworks have each been developed extensively, their convergence remains largely uncharted. In this paper, we present two novel constructs: *HyperEntropy*, which integrates entropy into a hyperstructural context, and *SuperHyperEntropy*, which further generalizes the idea to n-superhyperstructures. We establish their fundamental properties, illustrate them with concrete examples, and explore potential applications. Additionally, we present a preliminary example of their application in engineering. Our aim is to provide a foundation for future investigations that bridge these rich theoretical domains (see Table 2).

2. Preliminaries

This section provides an overview of the fundamental concepts and definitions essential for the discussions in this paper. Throughout this work, we consider only finite sets and finite structures, unless explicitly stated otherwise.

Table 2
Summary of Entropy, HyperEntropy, and SuperHyperEntropy

Concept	Domain	Definition	Interpretation
Entropy	Elements $x \in S$	$H(X) = -\sum_{x \in S} p(x) \log p(x)$	Measures the average unpredictability of a discrete random variable.
Hyper Entropy	Subsets $A \subseteq S$	$H_{\mathrm{h}}(P) = -\sum_{A \subseteq S} p(A) \log p(A)$	Quantifies uncertainty when outcomes are set-valued (subsets of S).
Super Hyper Entropy	Iterated powerset $P_n(S)$	$H_n(P) = -\sum_{A \in P_n(S)} p(A) \log p(A)$	Extends HyperEntropy to hierarchical, n-level set-valued outcomes.

2.1 Classical Structures, Hyperstructures, and n-Superhyperstructures

A *hyperstructure* generalizes a classical structure by permitting operations on subsets of the underlying set, producing subsets as outputs (cf. [(5)–(7), (17)]). Building on this, an *n-superhyperstructure* iterates the powerset construction n times, thereby capturing hierarchical or multi-level interactions among elements (cf. [(8), (9), (11), (18), (19)]). The formal definitions and representative examples are given below.

Definition 2.1 (Base Set)**.** *A base set S is any nonempty collection of elements under consideration. All subsequent constructions—such as power sets or iterated power sets—derive their members from S:*

$$S = \{\, x \mid x \text{ belongs to the universe of discourse} \}.$$

Definition 2.2 (Powerset)**.** *Given a set S, its* powerset, *denoted $\mathcal{P}(S)$, is the set of every subset of S, including the empty set and S itself:*

$$\mathcal{P}(S) = \{\, A \mid A \subseteq S \}.$$

Definition 2.3 (n-th Powerset)**.** *[(20)–(23)] The n-th powerset of S, written $\mathcal{P}_n(S)$, is defined by iteration:*

$$\mathcal{P}_1(S) = \mathcal{P}(S), \quad \mathcal{P}_{k+1}(S) = \mathcal{P}\big(P_k(S)\big) \quad (k \geq 1).$$

The nonempty version, $\mathcal{P}_n^*(S)$, omits the empty set at each stage:

$$\mathcal{P}_1^*(S) = \mathcal{P}(S) \setminus \{\emptyset\}, \quad \mathcal{P}_{k+1}^*(S) = \mathcal{P}\big(\mathcal{P}_k^*(S)\big) \setminus \{\emptyset\}.$$

Example 2.4 (Corporate Organization as a Third–Level Powerset)**.** *Consider a company with four employees:*

$$S = \{\textit{Yutaka},\ \textit{Hiroko},\ \textit{Tae},\ \textit{Shinya}\}.$$

- First–level powerset $P_1(S) = \mathcal{P}(S)$ *consists of all possible* project teams, *for example* $\{\textit{Yutaka}, \textit{Hiroko}\}$ *or* $\{\textit{Tae}, \textit{Shinya}\}$.
- Second–level powerset $P_2(S) = \mathcal{P}(P_1(S))$ *is the collection of* departments, *each of which is a set of project teams. For instance,*

  $$\big\{\{\textit{Yutaka}, \textit{Hiroko}\},\ \{\textit{Tae}\}\big\}$$

  might represent the "R&D" department.

- Third–level powerset $P_3(S) = \mathcal{P}(P_2(S))$ *organizes* divisions *as sets of departments. For example,*
$$\Big\{ \{\{\textit{Yutaka}, \textit{Hiroko}\}, \{\textit{Tae}\}\}, \{\{\textit{Hiroko}, \textit{Shinya}\}, \{\textit{Yutaka}, \textit{Tae}, \textit{Shinya}\}\} \Big\}$$
could represent the "Technology" division.

Thus the iterative construction
$$P_1(S) \to P_2(S) \to P_3(S)$$
captures the hierarchy: employees → teams → departments → divisions, illustrating a real–world application of the n-th powerset with $n = 3$.

Definition 2.5 (Classical Structure)**.** *A* Classical Structure *consists of a set H together with one or more operations $\# : H^m \to H$ ($m \geq 1$), satisfying axioms such as associativity or identity. Examples include groups $(H, \cdot)$, rings $(H, +, \times)$, and fields.*

Definition 2.6 (Hyperoperation)**.** *(cf.[(24)]) A* hyperoperation *on S replaces the codomain of a binary operation by the powerset of S:*
$$\circ : S \times S \longrightarrow \mathcal{P}(S).$$
Thus, combining two elements yields a (nonempty) subset of S rather than a single element.

Definition 2.7 (Hyperstructure)**.** *(cf.[(5), (25), (26)]) A* Hyperstructure *on a base set S is a pair $\big(\mathcal{P}(S), \circ\big)$, where $\circ$ is a hyperoperation on $\mathcal{P}(S)$. This allows one to apply "set-valued" operations to subsets of S.*

Example 2.8 (Chemical Reaction Hyperstructure)**.** *(cf.[(27)–(30)]) Let*
$$S = \{\mathrm{H_2},\ \mathrm{O_2},\ \mathrm{CO},\ \mathrm{N_2}\}.$$
Define a binary hyperoperation $\circ\colon S \times S \to \mathcal{P}(S)$ by
$$a \circ b \ = \ \{a, b\} \ \cup \begin{cases} \{\mathrm{H_2O}\}, & \textit{if } \{a, b\} = \{\mathrm{H_2}, \mathrm{O_2}\}, \\ \{\mathrm{CO_2}\}, & \textit{if } \{a, b\} = \{\mathrm{CO}, \mathrm{O_2}\}, \\ \emptyset, & \textit{otherwise}, \end{cases}$$
so that even non-reactive pairs still yield their reactants $\{a, b\}$. We then extend $\circ$ to a hyperoperation on $\mathcal{P}(S)$:
$$A \circ' B = \bigcup_{x \in A} \bigcup_{y \in B} (x \circ y), \qquad A, B \subseteq S.$$
Thus $\big(\mathcal{P}(S), \circ'\big)$ is a hyperstructure modeling chemical mixing.

Illustrative scenario: *Let*
$$A = \{\mathrm{H_2}, \mathrm{CO}\}, \quad B = \{\mathrm{O_2}\}.$$
Then
$$A \circ' B = \big\{\mathrm{H_2}, \mathrm{O_2}\big\} \ \cup \ \{\mathrm{H_2O}\} \ \cup \ \{\mathrm{CO}, \mathrm{O_2}\} \ \cup \ \{\mathrm{CO_2}\} = \{\mathrm{H_2},\ \mathrm{O_2},\ \mathrm{H_2O},\ \mathrm{CO},\ \mathrm{CO_2}\}.$$
This output captures all reactants and products formed when the sets of chemicals A and B are combined.

Definition 2.9 (SuperHyperOperation). *[(31)] Let H be nonempty and define iterated powersets $\mathcal{P}^k(H)$ as above. An (m,n)-*SuperHyperOperation *is an m-ary map*

$$\circ^{(m,n)} : H^m \longrightarrow \mathcal{P}^n(H),$$

where $\mathcal{P}^n(H)$ is the n-th powerset of H. When excluding the empty set one writes $\mathcal{P}^n_*(H)$.

Definition 2.10 (n-Superhyperstructure). *(cf.[(17), (31)–(33)]) An n-*Superhyperstructure *on S is the system*

$$\mathcal{SH}_n = \big(P_n(S), \circ\big),$$

where $\circ$ acts on elements of the n-th powerset $P_n(S)$. This framework captures nested, hierarchical operations through repeated powerset constructions.

Example 2.11 (Library Classification as a 2-Superhyperstructure). *Consider a library in which:*

$$S = \{\textit{all books in the library}\}.$$

- *The first–level powerset $P_1(S) = \mathcal{P}(S)$ is the collection of all genres, each genre being a subset of books.*
- *The second–level powerset $P_2(S) = \mathcal{P}\big(P_1(S)\big)$ is the set of all* collections of genres. *For example, $\{\{\textit{Fantasy}\}, \{\textit{Science Fiction}, \textit{Mystery}\}\}$ is one element of $P_2(S)$.*

Define a binary operation

$$\circ\colon P_2(S) \times P_2(S) \longrightarrow P_2(S)$$

by

$$X \circ Y = X \cup Y, \qquad X, Y \subseteq P_1(S).$$

Verification:

1. Closure: *Since the union of two collections of genres is again a collection of genres, $X \cup Y \in P_2(S)$.*
2. Associativity: *Union on any set is associative:* $(X \cup Y) \cup Z = X \cup (Y \cup Z)$.
3. Commutativity: $X \cup Y = Y \cup X$.
4. Identity (optional): *The empty collection $\emptyset \in P_2(S)$ serves as a neutral element:* $X \cup \emptyset = X$.

Thus $\big(P_2(S), \circ\big)$ is a concrete instance of a 2-Superhyperstructure, modeling how library genre collections can be merged.

Example 2.12 (Corporate Hierarchy as a 3-Superhyperstructure). *Consider an organization whose employees form the base set*

$$S = \{\textit{Yutaka}, \textit{Hiroko}, \textit{Tae}, \dots\}.$$

- *The first–level powerset $P_1(S) = \mathcal{P}(S)$ is the collection of all possible* teams, *each being a subset of employees.*
- *The second–level powerset $P_2(S) = \mathcal{P}(P_1(S))$ is the set of all* departments, *each comprising a set of teams.*
- *The third–level powerset $P_3(S) = \mathcal{P}(P_2(S))$ is the collection of* divisions, *each consisting of a set of departments.*

We define a binary operation

$$\circ\colon P_3(S) \times P_3(S) \longrightarrow P_3(S) \quad \textit{by} \quad X \circ Y = X \cup Y \quad (X, Y \subseteq P_2(S)).$$

Verification:

1. Closure: *The union of two divisions is again a set of departments, so* $X \cup Y \in P_3(S)$.
2. Associativity: *Union is associative:* $(X \cup Y) \cup Z = X \cup (Y \cup Z)$.
3. Commutativity: *Union is commutative:* $X \cup Y = Y \cup X$.
4. Identity: *The empty division* $\emptyset \in P_3(S)$ *serves as a neutral element:* $X \cup \emptyset = X$.

Hence $\big(P_3(S), \circ\big)$ *is a concrete example of a* 3*-superhyperstructure, modeling how corporate divisions can be merged.*

2.2 Shannon Entropy

Entropy quantifies the average information-theoretic unpredictability in a discrete random variable's probability distribution of outcomes (cf.[(1)–(4), (34), (35)]). The definition and a simple example are provided as follows.

Definition 2.13 (Shannon Entropy). *(cf.[(1)–(4), (34), (35)]) Let* X *be a discrete random variable taking values in a finite set* $\mathcal{X}$ *with probability mass function* $p_X(x) = \Pr(X = x)$. *The* entropy *of* X *is defined as*

$$H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x),$$

where the base of the logarithm determines the units (bits if $\log$ *is base 2, nats if base* e*).*

Example 2.14 (Fair Coin Toss). *Consider a fair coin toss, where* $X \in \{\textit{Heads}, \textit{Tails}\}$ *with* $\Pr(X = \textit{Heads}) = \Pr(X = \textit{Tails}) = 0.5$. *Then*

$$H(X) = -\big(0.5 \log_2 0.5 + 0.5 \log_2 0.5\big) = 1 \textit{ bit}.$$

This quantifies the maximum uncertainty when flipping a fair coin.

Example 2.15 (Three-State Weather Forecast). *A weather forecast is a scientific prediction of atmospheric conditions like temperature, precipitation, and wind for future time periods (cf.[(36), (37)]). Let* $X \in \{\textit{Sunny}, \textit{Cloudy}, \textit{Rainy}\}$ *with probabilities* $\Pr(X = \textit{Sunny}) = 0.7$, $\Pr(X = \textit{Cloudy}) = 0.2$, $\Pr(X = \textit{Rainy}) = 0.1$. *Then*

$$H(X) = -\big(0.7 \log_2 0.7 + 0.2 \log_2 0.2 + 0.1 \log_2 0.1\big) \approx 1.156 \textit{ bits}.$$

This measures the unpredictability in tomorrow's weather given these probabilities.

3. Result of This Paper

In this section, we present and explain the main results of this paper. Specifically, we investigate the mathematical properties of HyperEntropy and SuperHyperEntropy.

3.1 Shannon HyperEntropy

Hyperentropy quantifies the uncertainty of a probability distribution over subsets of a finite set, extending Shannon entropy to set-valued variables. The definition of HyperEntropy is given as follows.

Definition 3.1 (Hyperentropy). *Let S be a finite set and $\mathcal{P}(S)$ its powerset. Suppose*

$$p\colon \mathcal{P}(S) \;\longrightarrow\; [0,1]$$

is a probability mass function on subsets of S, i.e. $\sum_{A\subseteq S} p(A) = 1$. The hyperentropy *of p is defined by*

$$H_{\mathrm{h}}(p) \;=\; -\sum_{A\subseteq S} p(A)\,\log p(A),$$

where $\log$ denotes the natural logarithm (nats) or base-2 logarithm (bits).

Example 3.2 (Gourmet Pizza Topping Hyperentropy). *Let the base set of pizza toppings be*

$$S = \{\textit{Pepperoni},\ \textit{Mushrooms},\ \textit{Onions},\ \textit{Olives}\},$$

so that each subset $A \subseteq S$ corresponds to one possible topping combination. Suppose customer order frequencies yield the distribution

$$p(\{\textit{Pepperoni}\}) = 0.35, \quad p(\{\textit{Mushrooms}, \textit{Onions}\}) = 0.25,$$

$$p(\{\textit{Pepperoni}, \textit{Olives}\}) = 0.25, \quad p(\{\textit{Mushrooms}, \textit{Onions}, \textit{Olives}\}) = 0.15,$$

with $p(A) = 0$ for all other $A \subseteq S$. The hyperentropy of this distribution is

$$H_{\mathrm{h}}(p) = -\big(0.35\log 0.35 + 0.25\log 0.25 + 0.25\log 0.25 + 0.15\log 0.15\big) \;\approx\; 1.3452 \textit{ nats}.$$

This value quantifies the uncertainty in predicting which combination of toppings a customer will choose.

Example 3.3 (Market Basket Hyperentropy). *A market basket is a collection of products or items commonly purchased together by customers during a single shopping transaction (cf.[(38), (39)]). Let the base set of products be*

$$S = \{\textit{Bread},\ \textit{Milk},\ \textit{Eggs},\ \textit{Cheese}\},$$

so that $\mathcal{P}(S)$ is the set of all possible shopping baskets. Suppose customer purchase patterns follow the distribution

$$p(\{\textit{Bread}\}) = 0.4, \quad p(\{\textit{Milk}, \textit{Eggs}\}) = 0.3,$$

$$p(\{\textit{Bread}, \textit{Cheese}\}) = 0.2, \quad p(\{\textit{Milk}, \textit{Eggs}, \textit{Cheese}\}) = 0.1,$$

and $p(A) = 0$ for all other $A \subseteq S$. Then the hyperentropy of this distribution is

$$H_{\mathrm{h}}(p) = -\big(0.4\log 0.4 + 0.3\log 0.3 + 0.2\log 0.2 + 0.1\log 0.1\big) \approx 1.2799 \textit{ nats}.$$

This quantifies the uncertainty in predicting which combination of items a customer will purchase.

Theorem 3.4 (Generalization of Shannon Entropy). *Let $q\colon S \to [0,1]$ be a probability distribution on the singletons of S, and extend q to $\mathcal{P}(S)$ by*

$$p(A) = \begin{cases} q(x), & A = \{x\}, \\ 0, & |A| \neq 1. \end{cases}$$

Then the hyperentropy $H_{\mathrm{h}}(p)$ reduces to the classical Shannon entropy:

$$H_{\mathrm{h}}(p) = -\sum_{A \subseteq S} p(A)\ \log p(A) = -\sum_{x \in S} q(x)\ \log q(x) = H(q).$$

Hence H_{h} generalizes Shannon entropy.

Proof. Since $p(A) = 0$ whenever $|A| \neq 1$, the sum $\sum_{A \subseteq S} p(A) \log p(A)$ reduces to

$$\sum_{x \in S} p(\{x\}) \log p(\{x\}) = \sum_{x \in S} q(x) \log q(x).$$

The negative of this is exactly the Shannon entropy $H(q)$. □

Theorem 3.5 (Expression in a Hyperstructure). *Let $\mathcal{H} = (\mathcal{P}(S), \circ)$ be a hyperstructure as in Definition 1.2 of [(11)]. Then the hyperentropy H_{h} can be interpreted as the Shannon entropy of a random element of $\mathcal{H}$. Concretely, if X is a random variable taking values in $\mathcal{P}(S)$ with $\Pr(X = A) = p(A)$, then*

$$H_{\mathrm{h}}(p) = H\big(X\big) = -\sum_{A \subseteq S} \Pr(X = A)\ \log \Pr(X = A).$$

Proof. By definition of the hyperstructure $\mathcal{H}$, its underlying set is $\mathcal{P}(S)$. A probability distribution on $\mathcal{P}(S)$ makes X a discrete random variable on that set. The Shannon entropy of X is $H(X) = -\sum_{A \subseteq S} \Pr(X = A)\ \log \Pr(X = A)$, which coincides with $H_{\mathrm{h}}(p)$. □

Theorem 3.6 (Non-negativity). *For any distribution p on $\mathcal{P}(S)$,*

$$H_{\mathrm{h}}(p)\ \geq\ 0,$$

with equality if and only if there exists a subset $A^* \subseteq S$ such that $p(A^*) = 1$.

Proof. Since each term $-p(A) \log p(A)$ is nonnegative (defining $0 \log 0 = 0$), their sum is nonnegative. Moreover, $-x \log x = 0$ precisely when $x \in \{0, 1\}$. Thus $H_{\mathrm{h}}(p) = 0$ iff exactly one $p(A^*) = 1$ and all others vanish. □

Theorem 3.7 (Maximum Entropy). *Let $N = |S|$. Then*

$$H_{\mathrm{h}}(p)\ \leq\ \log\big|\mathcal{P}(S)\big|\ =\ N \log 2,$$

with equality if and only if p is the uniform distribution on $\mathcal{P}(S)$, i.e. $p(A) = 2^{-N}$ for all $A \subseteq S$.

Proof. The function $-x \log x$ is strictly concave on $[0, 1]$. By Jensen's inequality,

$$-\sum_{A} p(A) \log p(A)\ \leq\ -\Big(\sum_{A} p(A)\Big)\ \log\Big(\sum_{A} p(A)\Big) = \log |\mathcal{P}(S)|,$$

with equality exactly when all $p(A)$ are equal, i.e. uniform. □

Theorem 3.8 (Concavity). *Hyperentropy is a concave functional: for any two distributions p, q on $\mathcal{P}(S)$ and any $t \in [0,1]$,*

$$H_{\mathrm{h}}\big(tp + (1-t)q\big) \;\geq\; t\,H_{\mathrm{h}}(p) \;+\; (1-t)\,H_{\mathrm{h}}(q).$$

Proof. For each $A \subseteq S$, the function $f(x) = -x \log x$ is concave on $[0,1]$. Therefore, by pointwise application and summation,

$$\sum_{A} f\big(tp(A) + (1-t)q(A)\big) \;\geq\; t \sum_{A} f\big(p(A)\big) \;+\; (1-t) \sum_{A} f\big(q(A)\big),$$

which is precisely the desired inequality. □

Theorem 3.9 (Additivity for Independent Hypervariables). *Let S and T be disjoint finite sets, and let p on $\mathcal{P}(S)$, q on $\mathcal{P}(T)$ be two independent distributions. Define the joint distribution r on $\mathcal{P}(S \cup T)$ by*

$$r(A \cup B) \;=\; p(A)\,q(B), \quad A \subseteq S,\; B \subseteq T.$$

Then

$$H_{\mathrm{h}}(r) = H_{\mathrm{h}}(p) \;+\; H_{\mathrm{h}}(q).$$

Proof. By definition,

$$H_{\mathrm{h}}(r) = -\sum_{A \subseteq S} \sum_{B \subseteq T} p(A)q(B)\, \log\big(p(A)q(B)\big).$$

Since $\log(p(A)q(B)) = \log p(A) + \log q(B)$, the double sum splits:

$$-\sum_{A,B} p(A)q(B)\, \log p(A) \;-\; \sum_{A,B} p(A)q(B)\, \log q(B) = H_{\mathrm{h}}(p) + H_{\mathrm{h}}(q).$$

□

Theorem 3.10 (Data-Processing Inequality). *Let $f\colon \mathcal{P}(S) \to \mathcal{X}$ be any deterministic mapping into a finite set $\mathcal{X}$. Define the induced distribution r on $\mathcal{X}$ by*

$$r(x) = \sum_{A\,:\,f(A)=x} p(A).$$

Then

$$H(r) \;\leq\; H_{\mathrm{h}}(p),$$

where $H(r)$ is the ordinary Shannon entropy of r.

Proof. Partitioning the sum for $H_{\mathrm{h}}(p)$ according to preimages of f and applying Jensen's inequality on each block yields the inequality; see standard proofs of data-processing for discrete entropy. □

3.2 Shannon SuperHyperEntropy

Superhyperentropy measures the uncertainty of a probability distribution over n-th iterated powerset elements, extending set-valued entropy to multi-level hierarchical outcomes. The definition of SuperHyperEntropy is given as follows.

Definition 3.11 (n-Superhyperentropy). *Let $n \in \mathbb{N}$ and S a finite set. Write*

$$V^{(n)} = \mathcal{P}^n(S), \quad M_n = \left|V^{(n)}\right|.$$

*Let $p\colon V^{(n)} \to [0,1]$ be a probability mass function on $V^{(n)}$, so that $\sum_{A \in V^{(n)}} p(A) = 1$. The n-*superhyperentropy $H_n(p)$ *is*

$$H_n(p) = -\sum_{A \in V^{(n)}} p(A)\, \log p(A).$$

Example 3.12 (2-Superhyperentropy: Cross-Functional Committee Uncertainty). *Let the base set of project teams be*

$$S = \big\{\mathit{Team}_{AB},\, \mathit{Team}_{AC},\, \mathit{Team}_{BC}\big\},$$

where $\mathit{Team}_{AB} = \{A, B\}$, $\mathit{Team}_{AC} = \{A, C\}$, $\mathit{Team}_{BC} = \{B, C\}$. Then

$$\mathcal{P}^1(S) = \Big\{\{\mathit{Team}_{AB}\},\, \{\mathit{Team}_{AC}\},\, \{\mathit{Team}_{BC}\}\Big\},$$

and

$$\mathcal{P}^2(S) = \mathcal{P}\big(\mathcal{P}^1(S)\big)$$
$$= \Big\{\emptyset,\, \{\{\mathit{Team}_{AB}\}\},$$
$$\{\{\mathit{Team}_{AC}\}\},\, \{\{\mathit{Team}_{BC}\}\},\, \{\{\mathit{Team}_{AB}\}, \{\mathit{Team}_{AC}\}\},\, \dots\Big\}.$$

We focus on the three cross-functional committees of interest:

$$C_1 = \{\{\mathit{Team}_{AB}\}, \{\mathit{Team}_{BC}\}\},$$
$$C_2 = \{\{\mathit{Team}_{AB}\}, \{\mathit{Team}_{AC}\}\},$$
$$C_3 = \{\{\mathit{Team}_{AC}\}, \{\mathit{Team}_{BC}\}\}.$$

Define a probability mass function $p\colon \mathcal{P}^2(S) \to [0,1]$ by

$$p(C_1) = 0.5, \quad p(C_2) = 0.3, \quad p(C_3) = 0.2,$$

and $p(A) = 0$ for all other $A \in \mathcal{P}^2(S)$. Then the 2-superhyperentropy is

$$H_2(p) = -\sum_{i=1}^{3} p(C_i)\, \log p(C_i) = -\big(0.5 \log 0.5 + 0.3 \log 0.3 + 0.2 \log 0.2\big) \approx 1.0297 \text{ nats}.$$

This quantifies the uncertainty in which cross-functional committee of project teams will be formed.

Example 3.13 (2-Superhyperentropy: Salad Customization Uncertainty). *Let the base set of ingredients be*

$$S = \{\mathit{Lettuce},\, \mathit{Tomato},\, \mathit{Cheese}\}.$$

Then the set of all possible two-ingredient toppings is

$$\mathcal{P}^1(S) = \big\{\{\mathit{Lettuce}, \mathit{Tomato}\},\ \{\mathit{Lettuce}, \mathit{Cheese}\},\ \{\mathit{Tomato}, \mathit{Cheese}\}\big\},$$

which we denote by T_1, T_2, T_3. Next,

$$\mathcal{P}^2(S) = \mathcal{P}\big(\mathcal{P}^1(S)\big)$$

is the collection of all selections of topping pairs. Consider three popular salad combos:

$$C_A = \{T_1, T_2\}, \quad C_B = \{T_1, T_3\}, \quad C_C = \{T_2, T_3\}.$$

Define a probability mass function p on these combos by

$$p(C_A) = 0.6, \quad p(C_B) = 0.25, \quad p(C_C) = 0.15,$$

and $p(X) = 0$ for all other $X \in \mathcal{P}^2(S)$. Then the 2-superhyperentropy is

$$H_2(p) = -\sum_{X \in \{C_A, C_B, C_C\}} p(X)\, \log p(X) = -\big(0.6 \log 0.6 + 0.25 \log 0.25 + 0.15 \log 0.15\big) \approx 0.938 \text{ nats}.$$

This value quantifies the uncertainty in randomly choosing one of these three salad configurations.

Example 3.14 (3-Superhyperentropy: Product Configuration Uncertainty)**.** *A product configuration is a specific arrangement of components, features, or options selected to meet customer or production requirements (cf.[(40)–(42)]). Let the base set of modules be*

$$S = \{M_1,\ M_2,\ M_3\},$$

where $M_1 = \{Gear, Bolt\}$, $M_2 = \{Bolt, Panel\}$, $M_3 = \{Panel, Gear\}$. Then

$$\mathcal{P}^1(S) = \big\{\{M_1\}, \{M_2\}, \{M_3\}, \{M_1, M_2\}, \{M_2, M_3\}, \{M_1, M_3\}\big\},$$

the set of all possible subassemblies of one or two modules.

Next,

$$\mathcal{P}^2(S) = \mathcal{P}\big(\mathcal{P}^1(S)\big)$$

is the collection of all combinations of subassemblies. We focus on three subassemblies of interest:

$$SA_1 = \{M_1, M_2\}, \quad SA_2 = \{M_2, M_3\}, \quad SA_3 = \{M_1, M_3\}.$$

Finally,

$$\mathcal{P}^3(S) = \mathcal{P}\big(\mathcal{P}^2(S)\big)$$

is the set of all possible product configurations, each a collection of subassemblies. Consider three product configurations:

$$P_A = \{SA_1, SA_2\}, \quad P_B = \{SA_2, SA_3\}, \quad P_C = \{SA_1, SA_3\}.$$

Define a probability mass function $p\colon \mathcal{P}^3(S) \to [0, 1]$ by

$$p(P_A) = 0.5, \quad p(P_B) = 0.3, \quad p(P_C) = 0.2,$$

and $p(X) = 0$ for all other $X \in \mathcal{P}^3(S)$. The 3-superhyperentropy is then

$$H_3(p) = -\sum_{X \in \{P_A, P_B, P_C\}} p(X)\, \log p(X) = -\big(0.5 \log 0.5 + 0.3 \log 0.3 + 0.2 \log 0.2\big) \approx 1.0297 \text{ nats}.$$

This value quantifies the uncertainty in randomly selecting one of the three final product configurations.

Example 3.15 (3-Superhyperentropy: Corporate Division Uncertainty)**.** *Let the base set of employees be*

$$S = \{\textit{Yutaka},\ \textit{Hiroko},\ \textit{Tae}\}.$$

Then the set of all possible teams is

$$\mathcal{P}^1(S) = \big\{\{\textit{Yutaka}, \textit{Hiroko}\},\ \{\textit{Yutaka}, \textit{Tae}\},\ \{\textit{Hiroko}, \textit{Tae}\}\big\},$$

denoted $\textit{Team}_{AB}, \textit{Team}_{AC}, \textit{Team}_{BC}$.

Next, the collection of all departments (each a set of teams) is

$$\mathcal{P}^2(S) = \Big\{\{\textit{Team}_{AB}, \textit{Team}_{AC}\},\ \{\textit{Team}_{AB}, \textit{Team}_{BC}\},\ \{\textit{Team}_{AC}, \textit{Team}_{BC}\}\Big\},$$

denoted $\textit{Dept}_1, \textit{Dept}_2, \textit{Dept}_3$.

Finally, the set of all divisions (each a set of departments) is

$$\mathcal{P}^3(S) = \Big\{\{\textit{Dept}_1, \textit{Dept}_2\},\ \{\textit{Dept}_1, \textit{Dept}_3\},\ \{\textit{Dept}_2, \textit{Dept}_3\}\Big\},$$

denoted $\textit{Div}_A, \textit{Div}_B, \textit{Div}_C$.

Suppose the probability mass function on divisions is

$$p(\textit{Div}_A) = 0.4, \quad p(\textit{Div}_B) = 0.35, \quad p(\textit{Div}_C) = 0.25,$$

and $p(X) = 0$ *for all other* $X \in \mathcal{P}^3(S)$*. Then the 3-superhyperentropy is*

$$H_3(p) = -\sum_{X \in \{\textit{Div}_A, \textit{Div}_B, \textit{Div}_C\}} p(X)\ \log p(X)$$

$$= -\big(0.4 \log 0.4 + 0.35 \log 0.35 + 0.25 \log 0.25\big) \approx 1.0805\ \textit{nats}.$$

This quantifies the uncertainty in randomly selecting one of the three corporate divisions.

Example 3.16 (4-Superhyperentropy: Event Planning Uncertainty)**.** *In a large catering company, events are organized through four hierarchical levels of grouping:*

- ***Level 0 (*** S***):*** *Individual dishes* $S = \{\textit{Chicken Curry},\ \textit{Beef Stew},\ \textit{Grilled Fish}\}$.
- ***Level 1 (*** $P_1(S)$***):*** *Menus, each a subset of dishes, e.g.* $\{\textit{Chicken Curry}, \textit{Beef Stew}\}$ *or* $\{\textit{Grilled Fish}\}$.
- ***Level 2 (*** $P_2(S)$***):*** *Buffet packages, each a collection of menus, e.g.*

$$\{\{\textit{Chicken Curry}, \textit{Beef Stew}\}, \{\textit{Grilled Fish}\}\}.$$

- ***Level 3 (*** $P_3(S)$***):*** *Catering bundles, each a set of buffet packages, e.g.*

$$\{\textit{Package}_1, \textit{Package}_2\}.$$

- ***Level 4 (*** $P_4(S)$***):*** *Event plans, each grouping several catering bundles.*

Concretely, let

$$B_1 = \{\{\textit{Chicken Curry}, \textit{Beef Stew}\}, \{\textit{Grilled Fish}\}\},$$
$$B_2 = \{\{\textit{Chicken Curry}\}, \{\textit{Beef Stew}, \textit{Grilled Fish}\}\},$$
$$B_3 = \{\{\textit{Beef Stew}\}, \{\textit{Chicken Curry}, \textit{Grilled Fish}\}\},$$

$$C_1 = \{B_1, B_2\},$$
$$C_2 = \{B_1, B_3\},$$
$$C_3 = \{B_2, B_3\},$$

$$E_1 = \{C_1, C_2\},$$
$$E_2 = \{C_1, C_3\},$$
$$E_3 = \{C_2, C_3\}.$$

Here $E_1, E_2, E_3 \in P_4(S)$ represent three competing event plans.

Define a probability mass function $p\colon P_4(S) \to [0,1]$ by

$$p(E_1) = 0.5, \quad p(E_2) = 0.3, \quad p(E_3) = 0.2,$$

$$p(X) = 0 \textit{ for all other } X \in P_4(S).$$

Then the 4-superhyperentropy is

$$H_4(p) = -\sum_{i=1}^{3} p(E_i)\ \log p(E_i)$$

$$= -\big(0.5 \log 0.5 + 0.3 \log 0.3 + 0.2 \log 0.2\big)$$
$$\approx\ 1.0297 \textit{ nats}.$$

This value captures the uncertainty in selecting one of the three detailed event plans, each of which aggregates multiple bundles, packages, menus, and dishes.

Theorem 3.17 (Reduction to Shannon Entropy). *If $n = 0$, then $V^{(0)} = S$ and $H_0(p)$ is exactly the Shannon entropy of p on S:*

$$H_0(p) = -\sum_{x \in S} p(x)\ \log p(x).$$

Proof. By definition $\mathcal{P}^0(S) = S$, so the sum in $H_0(p)$ ranges over elements of S. Hence

$$H_0(p) = -\sum_{A \in S} p(A) \log p(A) = -\sum_{x \in S} p(x) \log p(x),$$

which is the classical Shannon entropy. □

Theorem 3.18 (Reduction to Hyperentropy). *If $n = 1$, then $V^{(1)} = \mathcal{P}(S)$ and $H_1(p)$ coincides with the hyperentropy defined on subsets of S:*

$$H_1(p)\ =\ -\sum_{A \subseteq S} p(A)\ \log p(A).$$

Proof. By the definition of n-superhyperentropy, for any $n \in \mathbb{N}$ and any probability mass function $p : V^{(n)} \to [0,1]$ with $\sum_{A \in V^{(n)}} p(A) = 1$, we have

$$H_n(p) \;=\; -\sum_{A \in V^{(n)}} p(A)\, \log p(A).$$

Specializing to $n = 1$, observe that

$$V^{(1)} = \mathcal{P}^1(S) = \mathcal{P}(S),$$

so the defining sum becomes

$$H_1(p) = -\sum_{A \in \mathcal{P}(S)} p(A)\, \log p(A)$$

$$= -\sum_{A \subseteq S} p(A)\, \log p(A).$$

By Definition of hyperentropy on $\mathcal{P}(S)$, this expression is exactly $H_{\mathrm{h}}(p)$. Hence $H_1(p) = H_{\mathrm{h}}(p)$, as claimed. □

Theorem 3.19 (Expression in an n-Superhyperstructure)**.** *Let*

$$\mathcal{SH}_n = \big(\mathcal{P}^n(S),\, \circ\big)$$

be an n-superhyperstructure. If X is a random element of $\mathcal{SH}_n$ with $\Pr(X = A) = p(A)$ for each $A \in \mathcal{P}^n(S)$, then

$$H_n(p) = H(X) = -\sum_{A \in \mathcal{P}^n(S)} \Pr(X = A)\, \log \Pr(X = A).$$

Thus n-superhyperentropy is the Shannon entropy of X in the superhyperstructure.

Proof. By construction, the underlying set of $\mathcal{SH}_n$ is $\mathcal{P}^n(S)$. A distribution p on this set makes X a discrete random variable. Its Shannon entropy is

$$H(X) = -\sum_{A \in \mathcal{P}^n(S)} \Pr(X = A)\, \log \Pr(X = A) = -\sum_{A \in V^{(n)}} p(A) \log p(A) = H_n(p).$$

□

Theorem 3.20 (Non-negativity)**.** *For any probability mass function $p\colon V^{(n)} \to [0,1]$,*

$$H_n(p) \;\geq\; 0,$$

with equality if and only if $p(A^*) = 1$ for some $A^* \in V^{(n)}$.

Proof. Each term $-p(A) \log p(A)$ is nonnegative (defining $0 \log 0 = 0$). Hence the sum is nonnegative. Equality holds exactly when one probability is 1 and the rest are 0. □

Theorem 3.21 (Maximum Entropy)**.**

$$H_n(p) \;\leq\; \log M_n,$$

with equality if and only if p is the uniform distribution on $V^{(n)}$, i.e. $p(A) = 1/M_n$ for all A.

Proof. The function $-x \log x$ is strictly concave on $[0, 1]$. By Jensen's inequality,

$$H_n(p) = \sum_A -p(A) \log p(A) \leq -\Big(\sum_A p(A)\Big) \log\Big(\sum_A p(A)\Big) = \log M_n,$$

with equality exactly when all $p(A)$ are equal. □

Theorem 3.22 (Concavity). *If p, q are two distributions on $V^{(n)}$ and $t \in [0, 1]$, then*

$$H_n\big(tp + (1-t)q\big) \;\geq\; t\, H_n(p) \;+\; (1-t)\, H_n(q).$$

Proof. Define the function

$$f(x) = \begin{cases} -x \log x, & x > 0, \\ 0, & x = 0, \end{cases}$$

on the interval $[0, 1]$. A direct computation shows

$$f''(x) = -\frac{1}{x} < 0 \quad \text{for } x > 0,$$

so f is strictly concave on $(0, 1]$ (and continuous at 0). Now let p, q be two probability mass functions on $V^{(n)}$ and fix $t \in [0, 1]$. For each $A \in V^{(n)}$, concavity of f gives

$$f\big(t\, p(A) + (1-t)\, q(A)\big) \;\geq\; t\, f\big(p(A)\big) \;+\; (1-t)\, f\big(q(A)\big).$$

Summing this inequality over all $A \in V^{(n)}$ yields

$$H_n\big(tp + (1-t)q\big) = \sum_{A \in V^{(n)}} f\big(t\, p(A) + (1-t)\, q(A)\big) \;\geq$$

$$t \sum_A f\big(p(A)\big) \;+\; (1-t) \sum_A f\big(q(A)\big)$$

$$= t\, H_n(p) + (1-t)\, H_n(q).$$

This establishes the concavity of H_n. □

Theorem 3.23 (Chain Rule). *Define the marginal on $V^{(n-1)}$,*

$$p^{(n-1)}(B) = \sum_{A \supseteq B,\, A \in V^{(n)}} p(A),$$

$$B \in V^{(n-1)},$$

and the conditional

$$p(A \mid B) = \frac{p(A)}{p^{(n-1)}(B)}$$

when $A \supseteq B$. Then

$$H_n(p) = H_{n-1}\big(p^{(n-1)}\big)$$

$$+ \sum_{B \in V^{(n-1)}} p^{(n-1)}(B)\, H_1\big(p(\cdot \mid B)\big).$$

Proof. Split the sum defining $H_n(p)$ according to the level- $(n-1)$ subset $B \subseteq S$:

$$H_n(p) = -\sum_{B \in V^{(n-1)}} \sum_{A \supseteq B} p(A) \log p(A).$$

Write $p(A) = p^{(n-1)}(B)\, p(A \mid B)$ and expand:

$$\begin{aligned} -\sum_{B} \sum_{A \supseteq B} p(A) \log\big[p^{(n-1)}(B) p(A \mid B)\big] &= -\sum_{B} p^{(n-1)}(B) \log p^{(n-1)}(B) \\ &\quad - \sum_{B} p^{(n-1)}(B) \sum_{A \supseteq B} p(A \mid B) \log p(A \mid B) \\ &= H_{n-1}\big(p^{(n-1)}\big) + \sum_{B} p^{(n-1)}(B)\, H_1\big(p(\cdot \mid B)\big). \end{aligned}$$

□

Theorem 3.24 (Data-Processing Inequality). *Let $f\colon V^{(n)} \to \mathcal{X}$ be a deterministic map into a finite set $\mathcal{X}$. Define the induced distribution $r\colon \mathcal{X} \to [0,1]$ by*

$$r(x) \;=\; \sum_{\substack{A \in V^{(n)} \\ f(A) = x}} p(A).$$

Then

$$H(r) \;\le\; H_n(p),$$

where $H(r) = -\sum_{x \in \mathcal{X}} r(x) \log r(x)$ is the Shannon entropy of r and $H_n(p) = -\sum_{A \in V^{(n)}} p(A) \log p(A)$.

Proof. Partition the sum defining $H_n(p)$ according to the fibers of f:

$$H_n(p) = -\sum_{A \in V^{(n)}} p(A)\, \log p(A) = -\sum_{x \in \mathcal{X}} \sum_{A : f(A) = x} p(A)\, \log p(A).$$

Fix $x \in \mathcal{X}$ and let

$$r(x) \;=\; \sum_{A : f(A) = x} p(A).$$

Define weights $\alpha_A = p(A)/r(x)$ for those A with $f(A) = x$. Then

$$\sum_{A : f(A) = x} p(A)\, \log p(A) = r(x) \sum_{A : f(A) = x} \alpha_A \log\big(r(x)\, \alpha_A\big) = r(x) \log r(x) + r(x) \sum_{A : f(A) = x} \alpha_A \log \alpha_A.$$

Since $\alpha_A \in [0,1]$ and $\sum \alpha_A = 1$, we have $\sum \alpha_A \log \alpha_A \le 0$. Hence

$$\sum_{A : f(A) = x} p(A)\, \log p(A) \;\le\; r(x)\, \log r(x).$$

Multiplying by -1 and summing over x gives

$$H_n(p) = -\sum_{x} \sum_{A : f(A) = x} p(A)\, \log p(A) \;\ge\; -\sum_{x} r(x)\, \log r(x) = H(r),$$

as required. □

Theorem 3.25 (Permutation Invariance). *Let $\sigma : V^{(n)} \to V^{(n)}$ be any bijection, and define $p_\sigma(A) = p\big(\sigma(A)\big)$. Then*

$$H_n(p_\sigma) = H_n(p).$$

Proof.

$$H_n(p_\sigma) = -\sum_{A\in V^{(n)}} p_\sigma(A)\,\log p_\sigma(A) = -\sum_{A\in V^{(n)}} p\big(\sigma(A)\big)\,\log p\big(\sigma(A)\big).$$

Since σ is a bijection, reindexing $B = \sigma(A)$ shows this equals $-\sum_{B\in V^{(n)}} p(B)\log p(B) = H_n(p)$. □

Theorem 3.26 (Support Bound). *Let $\operatorname{supp}(p) = \{A \in V^{(n)} : p(A) > 0\}$ have size s. Then*

$$H_n(p) \;\le\; \log s,$$

with equality if and only if p is uniform on its support.

Proof. View p as a distribution on a set of size s. By the maximum-entropy principle,

$$H_n(p) \;\le\; \log s,$$

with equality precisely when $p(A) = 1/s$ for all $A \in \operatorname{supp}(p)$. □

Theorem 3.27 (Lower Bound via Maximal Probability). *Let $p_{\max} = \max_{A\in V^{(n)}} p(A)$. Then*

$$H_n(p) \;\ge\; -\log p_{\max}.$$

Proof. Since $p(A) \le p_{\max}$ for all A, we have $-\log p(A) \ge -\log p_{\max}$. Hence

$$H_n(p) = -\sum_A p(A)\log p(A) \;\ge\; -\sum_A p(A)\,\log p_{\max} = -\log p_{\max}\sum_A p(A) = -\log p_{\max}.$$

□

Theorem 3.28 (Kullback–Leibler Divergence to Uniform). *Let $u(A) = 1/M_n$ be the uniform distribution on $V^{(n)}$. Then the Kullback–Leibler divergence*

$$D\big(p\|u\big) = \sum_{A\in V^{(n)}} p(A)\,\log\frac{p(A)}{u(A)} = \log M_n \;-\; H_n(p) \;\ge\; 0,$$

with equality if and only if p is uniform.

Proof. By definition,

$$D(p\|u) = \sum_A p(A)\big(\log p(A) + \log M_n\big) = \log M_n - \sum_A p(A)\log p(A) = \log M_n - H_n(p).$$

Nonnegativity of $D(p\|u)$ then yields $H_n(p) \le \log M_n$. □

Theorem 3.29 (Strict Concavity). *For any two distinct distributions $p \ne q$ on $V^{(n)}$ and any $t \in (0,1)$,*

$$H_n\big(tp + (1-t)q\big) \;>\; t\,H_n(p) \;+\; (1-t)\,H_n(q).$$

Proof. The function $f(x) = -x\log x$ is strictly concave on $(0,1)$. Hence for each A,

$$f\big(tp(A) + (1-t)q(A)\big) > t\,f\big(p(A)\big) + (1-t)\,f\big(q(A)\big).$$

Summing over all $A \in V^{(n)}$ yields the strict inequality for H_n. □

4. Additional Real-Life Example in Engineering

In this section, we examine an additional real-life example from the field of engineering.

Example 4.1 (4-Superhyperentropy: Hierarchical Control Architecture in Smart Manufacturing)**.** *Consider a smart manufacturing facility composed of modular units, each capable of autonomous operation. We model the uncertainty over system-level control plans using the notion of 4-superhyperentropy.*

Let the base set of atomic control modules *be:*

$$S = \{M_1, M_2, M_3, M_4\},$$

where:

- M_1*: Robotic arm controller,*
- M_2*: Temperature regulation module,*
- M_3*: Quality inspection unit,*
- M_4*: Conveyor belt control.*

Level 1 ($P_1(S) = \mathcal{P}(S)$***):*** *All subsets of modules represent* basic operational configurations.

Let us fix:

$$C_1 = \{M_1, M_2\}, \quad C_2 = \{M_2, M_3\},$$
$$C_3 = \{M_1, M_4\}, \quad C_4 = \{M_3, M_4\}.$$

Level 2 ($P_2(S)$***):*** *Each subset of* $\{C_1, C_2, C_3, C_4\}$ *is a* workflow plan, *consisting of selected configurations.*

Let us define:

$$W_1 = \{C_1, C_2\}, \quad W_2 = \{C_2, C_3\}, \quad W_3 = \{C_3, C_4\}.$$

Level 3 ($P_3(S)$***):*** *Subsets of* $\{W_1, W_2, W_3\}$ *define* production schemes, *grouping workflows into intermediate manufacturing strategies.*

Define:

$$P_1 = \{W_1, W_2\}, \quad P_2 = \{W_2, W_3\}, \quad P_3 = \{W_1, W_3\}.$$

Level 4 ($P_4(S)$***):*** *Subsets of* $\{P_1, P_2, P_3\}$ *represent* global production policies, *each being a system-wide plan across all units.*

Let:

$$G_1 = \{P_1, P_2\}, \quad G_2 = \{P_2, P_3\}, \quad G_3 = \{P_1, P_3\}.$$

Let $p : \mathcal{P}^4(S) \to [0, 1]$ *be a probability mass function modeling historical or predicted system usage:*

$$p(G_1) = 0.4, \quad p(G_2) = 0.35, \quad p(G_3) = 0.25,$$

and $p(X) = 0$ *for all other* $X \in \mathcal{P}^4(S)$*.*

Then the 4*-superhyperentropy is:*

$$H_4(p) = -\sum_{i=1}^{3} p(G_i) \log p(G_i) = -\left(0.4 \log 0.4 + 0.35 \log 0.35 + 0.25 \log 0.25\right) \approx 1.0805 \text{ nats}.$$

This entropy quantifies the uncertainty over which complete system-level production policy will be selected. Since the underlying objects are elements of $\mathcal{P}^4(S)$*, this models control complexity over four hierarchical engineering layers, making it useful for stochastic planning and robust optimization in cyber-physical production systems.*

Example 4.2 (4-Superhyperentropy in Engineering: Autonomous Vehicle Sensor Fusion)**.** *We consider an autonomous vehicle (AV) system equipped with multiple sensor types. These sensors operate under layered configurations, fusion mechanisms, and environmental conditions, forming a hierarchical decision framework. To assess the uncertainty in selecting a system-wide deployment strategy, we utilize 4-superhyperentropy.*

Level 0 (Base set S***):*** *Let the base set of sensors be*

$$S = \{\mathrm{L}, \mathrm{C}, \mathrm{R}, \mathrm{U}\},$$

with:

- L*: LiDAR,*
- C*: Camera,*
- R*: Radar,*
- U*: Ultrasonic sensor.*

Level 1 ($V^{(1)} = \mathcal{P}(S)$***):*** *Consider specific sensor configurations:*

$$\begin{aligned} S_1 &= \{\mathrm{L}, \mathrm{C}\}, \\ S_2 &= \{\mathrm{R}, \mathrm{U}\}, \\ S_3 &= \{\mathrm{C}, \mathrm{U}\}, \\ S_4 &= \{\mathrm{L}, \mathrm{R}\}. \end{aligned}$$

Level 2 ($V^{(2)} = \mathcal{P}(V^{(1)})$***):*** *Define fusion modes that combine sensor configurations:*

$$\begin{aligned} F_1 &= \{S_1, S_2\}, \\ F_2 &= \{S_3, S_4\}, \\ F_3 &= \{S_2, S_3\}. \end{aligned}$$

Level 3 ($V^{(3)} = \mathcal{P}(V^{(2)})$***):*** *Define operational contexts:*

$$\begin{aligned} C_1 &= \{F_1, F_2\}, \\ C_2 &= \{F_2, F_3\}, \\ C_3 &= \{F_1, F_3\}. \end{aligned}$$

Level 4 ($V^{(4)} = \mathcal{P}(V^{(3)})$***):*** *Define deployment strategies:*

$$\begin{aligned} D_1 &= \{C_1, C_3\}, \\ D_2 &= \{C_1, C_2\}, \\ D_3 &= \{C_2, C_3\}. \end{aligned}$$

Probability distribution: *Let the probability mass function* $p : V^{(4)} \to [0, 1]$ *be given by*

$$p(D_1) = 0.45, \quad p(D_2) = 0.30, \quad p(D_3) = 0.25,$$

and $p(X) = 0$ *for all other* $X \in V^{(4)} \setminus \{D_1, D_2, D_3\}$*.*

Computation of 4-superhyperentropy: *The 4-superhyperentropy is*

$$H_4(p) = -\sum_{i=1}^{3} p(D_i)\log p(D_i) = -\left(0.45\log 0.45 + 0.30\log 0.30 + 0.25\log 0.25\right) \approx 1.0619 \textbf{ nats}.$$

The value $H_4(p)$ reflects the degree of uncertainty in selecting a comprehensive deployment strategy D_i, each comprising specific combinations of operational contexts, fusion modes, and sensor configurations. This hierarchical entropy metric provides a principled way for AV engineers to quantify systemic uncertainty and assess the robustness of adaptive strategies under varying external factors (e.g., weather, traffic, sensor faults). Such rigorous modeling is indispensable in engineering reliable autonomous systems.

5. Conclusion and Future Works

In this paper, we introduced two novel constructs: *HyperEntropy*, which extends Shannon entropy to probability distributions over subsets of a base set, and *SuperHyperEntropy*, which further generalizes this concept to distributions defined on iterated powerset levels. These frameworks enable the quantification of uncertainty in hierarchical and set-valued contexts, offering a deeper and more flexible perspective within entropy theory.

In future work, we aim to explore extended entropy models that incorporate additional uncertainty frameworks such as Fuzzy Sets [(43)], Intuitionistic Fuzzy Sets [(44)], Neutrosophic Sets [(45)], Neutrosophic Offsets [(46)], Soft Sets [(47)], Hypersoft Sets [(48), (49)], Rough Sets [(50), (51)], Picture Fuzzy Sets[(52)], Hesitant Fuzzy Sets [(53), (54)], Quadri-Partitioned Neutrosophic Sets [(55)], HyperNeutrosophic Sets [(56), (57)], and Plithogenic Sets [(58), (59)]. These directions hold promise for further generalizing entropy in ways that reflect the complexity and indeterminacy present in real-world data and decision-making environments.

Acknowledgments

This research was not funded by any grant.

Conflict of Interest

The authors declare no conflicts of interest.

References

[1] Lin, J. (2002). Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory*, *37*(1), 145–151. https://doi.org/10.1109/18.61115

[2] Eskov, V., Eskov, V., Vochmina, Y. V., Gorbunov, D., & Ilyashenko, L. (2017). Shannon entropy in the research on stationary regimes and the evolution of complexity. *Moscow university physics bulletin*, *72*, 309–317. https://doi.org/https://doi.org/10.3103/S0027134917030067

[3] Wu, Y., Zhou, Y., Saveriades, G., Agaian, S., Noonan, J. P., & Natarajan, P. (2013). Local shannon entropy measure with statistical tests for image randomness. *Information Sciences*, *222*, 323–342. https://doi.org/https://doi.org/10.1016/j.ins.2012.07.049

[4] Tiwari, V. K., Jain, P. K., & Tandon, P. (2017). An integrated shannon entropy and topsis for product design concept evaluation based on bijective soft set. *Journal of Intelligent Manufacturing*, *30*, 1645–1658. https://doi.org/https://doi.org/10.1007/s10845-017-1346-y

[5] Corsini, P., & Leoreanu, V. (2013). *Applications of hyperstructure theory* (Vol. 5). Springer Science & Business Media.

[6] Oguz, G., & Davvaz, B. (2021). Soft topological hyperstructure. *J. Intell. Fuzzy Syst.*, *40*, 8755–8764. https://api.semanticscholar.org/CorpusID:234193021

[7] Adebisi, S. A., & Ajuebishi, A. P. (2025). The order involving the neutrosophic hyperstructures, the construction and setting up of a typical neutrosophic group. *HyperSoft Set Methods in Engineering*, *3*, 26–31. https://doi.org/https://doi.org/10.61356/j.hsse.2025.3426

[8] Al-Odhari, A. (2025). A brief comparative study on hyperstructure, super hyperstructure, and n-super superhyperstructure. *Neutrosophic Knowledge*, *6*, 38–49. https://doi.org/https://doi.org/10.5281/zenodo.15107294

[9] Fujita, T. (2025a). Antihyperstructure, neutrohyperstructure, and superhyperstructure. *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond*, 311.

[10] Das, A. K., Das, R., Das, S., Debnath, B. K., Granados, C., Shil, B., & Das, R. (2025). A comprehensive study of neutrosophic superhyper bci-semigroups and their algebraic significance. *Transactions on Fuzzy Sets and Systems*, *8*(2), 80. https://doi.org/10.71602/tfss.2025.1198050

[11] Smarandache, F. (2024a). Foundation of superhyperstructure & neutrosophic superhyperstructure. *Neutrosophic Sets and Systems*, *63*(1), 21. https://fs.unm.edu/nss8/index.php/111/article/view/3896

[12] Smarandache, F. (2020). *Extension of hypergraph to n-superhypergraph and to plithogenic n-superhypergraph, and extension of hyperalgebra to n-ary (classical-/neutro-/anti-) hyperalgebra*. Infinite Study.

[13] Alqahtani, M. (2025). Intuitionistic fuzzy quasi-supergraph integration for social network decision making. *International Journal of Analysis and Applications*, *23*, 137–137. https://doi.org/https://doi.org/10.28924/2291-8639-23-2025-137

[14] Fujita, T., & Smarandache, F. (2024). A concise study of some superhypergraph classes. *Neutrosophic Sets and Systems*, *77*, 548–593. https://fs.unm.edu/nss8/index.php/111/article/view/5416

[15] Fujita, T., & Ghaib, A. A. (2025). Toward a unified theory of brain hypergraphs and symptom hypernetworks in medicine and neuroscience. *Advances in Research*, *26*(3), 522–565. https://doi.org/https://doi.org/10.9734/air/2025/v26i31368

[16] Smarandache, F. (2022). *Introduction to superhyperalgebra and neutrosophic superhyperalgebra*. Infinite Study.

[17] Fujita, T. (2025b). An introduction and reexamination of hyperprobability and superhyperprobability: Comprehensive overview. *Asian Journal of Probability and Statistics*, *27*(5), 82–109. https://doi.org/https://doi.org/10.9734/ajpas/2025/v27i5757

[18] Fujita, T. (2024a). Rethinking strategic perception: Foundations and advancements in hypergame theory and superhypergame theory. *Prospects for Applied Mathematics and Data Analysis*, *4*(2), 01–14. https://doi.org/10.54216/PAMDA.040201

[19] Fujita, T. (2024b). Hyperalgorithms & superhyperalgorithms: A unified framework for higher-order computation. *Prospects for Applied Mathematics and Data Analysis*, *4*(1), 36–49. https://doi.org/10.54216/PAMDA.040104

[20] Smarandache, F. (2023). *Superhyperfunction, superhyperstructure, neutrosophic superhyperfunction and neutrosophic superhyperstructure: Current understanding and future directions*. Infinite Study.

[21] Fujita, T. (2025c). A theoretical investigation of quantum n-superhypergraph states. *Neutrosophic Optimization and Intelligent Systems*, *6*, 15–25. https://doi.org/10.61356/j.nois.2025.6544

[22] Fujita, T. (2025d). *Advancing uncertain combinatorics through graphization, hyperization, and uncertainization: Fuzzy, neutrosophic, soft, rough, and beyond*. Biblio Publishing.

[23] Smarandache, F. (2024b). The cardinal of the m-powerset of a set of n elements used in the super-hyperstructures and neutrosophic superhyperstructures. *Systems Assessment and Engineering Management*, *2*, 19–22. https://doi.org/https://doi.org/10.61356/j.saem.2024.2436

[24] Vougioukli, S. (2020). Helix hyperoperation in teaching research. *Science & Philosophy*, *8*(2), 157–163.

[25] Vougiouklis, T. (2015). Hypermathematics, hv-structures, hypernumbers, hypermatrices and lie-santilli admissibility. *American Journal of Modern Physics*, *4*(5), 38–46. https://doi.org/10.11648/j.ajmp.s.2015040501.15

[26] Ruggero, M. S., & Vougiouklis, T. (2017). Hyperstructures in lie-santilli admissibility and iso-theories. *Ratio Mathematica*, *33*, 151. https://doi.org/http://dx.doi.org/10.23755/rm.v33i0.374

[27] Al Tahan, M., & Davvaz, B. (2018). Weak chemical hyperstructures associated to electrochemical cells. *Iranian Journal of Mathematical Chemistry*, *9*(1), 65–75. https://doi.org/10.22052/ijmc.2017.88790.1294

[28] Davvaz, B. (2016). Weak algebraic hyperstructures as a model for interpretation of chemical reactions. *Iranian Journal of Mathematical Chemistry*, *7*(2), 267–283. https://doi.org/10.22052/ijmc.2016.13975

[29] Fujita, T. (2025e). Chemical hyperstructures, superhyperstructures, and shv-structures: Toward a generalized framework for hierarchical chemical modeling. *ChemRxiv*. https://doi.org/10.26434/chemrxiv-2025-zk8pm

[30] Al-Tahan, M., & Davvaz, B. (2022). Chemical hyperstructures for elements with four oxidation states. *Iranian Journal of Mathematical Chemistry*, *13*(2), 85–97. https://doi.org/10.22052/ijmc.2022.246174.1615

[31] Smarandache, F. (2024c). Superhyperstructure & neutrosophic superhyperstructure [Accessed: 2024-12-01]. https://fs.unm.edu/SHS/

[32] Jdid, M., Smarandache, F., & Fujita, T. (2025). A linear mathematical model of the vocational training problem in a company using neutrosophic logic, hyperfunctions, and superhyperfunction. *Neutrosophic Sets and Systems*, *87*, 1–11. https://doi.org/10.5281/zenodo.15558388

[33] Kargın, A., & Şahin, M. (2023). Superhyper groups and neutro–superhyper groups. *2023 Neutrosophic SuperHyperAlgebra And New Types of Topologies*, *25*.

[34] Hammer, D., Romashchenko, A., Shen, A., & Vereshchagin, N. (2000). Inequalities for shannon entropy and kolmogorov complexity. *Journal of Computer and System Sciences*, *60*(2), 442–464. https://doi.org/https://doi.org/10.1006/jcss.1999.1677

[35] Noorizadeh, S., & Shakerzadeh, E. (2010). Shannon entropy as a new measure of aromaticity, shannon aromaticity. *Physical Chemistry Chemical Physics*, *12*(18), 4742–4749. https://doi.org/10.1039/b916509f

[36] Riordan, D., & Hansen, B. K. (2002). A fuzzy case-based system for weather prediction. *Engineering Intelligent Systems for Electrical Engineering and Communications*, *10*(3), 139–146.

[37] Al-Matarneh, L., Sheta, A., Bani-Ahmad, S., Alshaer, J., & Al-Oqily, I. (2014). Development of temperature-based weather forecasting models using neural networks and fuzzy logic. *International journal of multimedia and ubiquitous engineering*, *9*(12), 343–366. https://doi.org/10.14257/ijmue.2014.9.12.31

[38] Valle, M. A., Ruz, G. A., & Morrás, R. (2018). Market basket analysis: Complementing association rules with minimum spanning trees. *Expert Systems with Applications*, *97*, 146–162.

[39] Raeder, T., & Chawla, N. V. (2011). Market basket analysis with networks. *Social network analysis and mining*, *1*, 97–113. https://doi.org/10.1007/s13278-010-0003-7

[40] Zhang, L. L. (2014). Product configuration: A review of the state-of-the-art and future research. *International Journal of Production Research*, *52*(21), 6381–6398. https://doi.org/10.1080/00207543.2014.942012

[41] Blecker, T. (2006). Product configuration systems. In *Encyclopedia of e-commerce, e-government, and mobile commerce* (pp. 941–947). IGI Global.
[42] Jiao, Y., & Yang, Y. (2019). A product configuration approach based on online data. *Journal of Intelligent Manufacturing*, *30*, 2473–2487. https://doi.org/https://doi.org/10.1007/s10845-018-1406-y
[43] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338–353.
[44] Atanassov, K. (2016). Intuitionistic fuzzy sets. *International journal bioautomation*, *20*, 1.
[45] Smarandache, F., & Jdid, M. (2023). An overview of neutrosophic and plithogenic theories and applications.
[46] Smarandache, F. (2016). Neutrosophic overset. *Neutrosophic Underset, and Neutrosophic Offset: Similarly for Neutrosophic Over-/Under-/Off-Logic, Probability, and Statistics*, *Pons Editions Brussels*.
[47] Maji, P. K., Biswas, R., & Roy, A. R. (2003). Soft set theory. *Computers & mathematics with applications*, *45*(4-5), 555–562.
[48] Smarandache, F. (2018). Extension of soft set to hypersoft set, and then to plithogenic hypersoft set. *Neutrosophic sets and systems*, *22*(1), 168–170. https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1309&context=math_fsp
[49] Musa, S. Y., & Asaad, B. A. (2022). Topological structures via bipolar hypersoft sets. *Journal of Mathematics*, *2022*(1), 2896053. https://doi.org/https://doi.org/10.1155/2022/2896053
[50] Pawlak, Z. (1982). Rough sets. *International journal of computer & information sciences*, *11*, 341–356. https://doi.org/https://doi.org/10.1007/BF01001956
[51] Pawlak, Z. (2012). *Rough sets: Theoretical aspects of reasoning about data* (Vol. 9). Springer Science & Business Media.
[52] Cuong, B. C., & Kreinovich, V. (2013). Picture fuzzy sets-a new concept for computational intelligence problems. *2013 third world congress on information and communication technologies (WICT 2013)*, 1–6. https://doi.org/10.1109/WICT.2013.7113099
[53] Torra, V., & Narukawa, Y. (2009). On hesitant fuzzy sets and decision. *2009 IEEE international conference on fuzzy systems*, 1378–1382. https://doi.org/10.1109/FUZZY.2009.5276884
[54] Beg, I., & Rashid, T. (2014). Hesitant intuitionistic fuzzy linguistic term sets. https://ifigenia.org/images/9/93/NIFS-20-3-53-64.pdf
[55] Hussain, S., Hussain, J., Rosyida, I., & Broumi, S. (2022). Quadripartitioned neutrosophic soft graphs. In *Handbook of research on advances and applications of fuzzy sets and logic* (pp. 771–795). IGI Global.
[56] Smarandache, F. (2017). *Hyperuncertain, superuncertain, and superhyperuncertain sets/logics/probabilities/statistics*. Infinite Study.
[57] Fujita, T., & Smarandache, F. (2025). A concise introduction to hyperfuzzy, hyperneutrosophic, hyperplithogenic, hypersoft, and hyperrough sets with practical examples. *Neutrosophic Sets and Systems*, *80*, 609–631. https://doi.org/10.5281/zenodo.14759385
[58] Abualhomos, M., Shihadeh, A., A Abubaker, A., Al-Husban, K., Fujita, T., Alsaraireh, A. A., Shatnawi, M., & Al-Husban, A. (2025). Unified framework for type-n extensions of fuzzy, neutrosophic, and plithogenic offsets: Definitions and interconnections. *Journal of Fuzzy Extension and Applications*. https://doi.org/https://doi.org/10.22105/jfea.2025.514314.1858
[59] Sultana, F., Gulistan, M., Ali, M., Yaqoob, N., Khan, M., Rashid, T., & Ahmed, T. (2023). A study of plithogenic graphs: Applications in spreading coronavirus disease (covid-19) globally. *Journal of ambient intelligence and humanized computing*, *14*(10), 13139–13159. https://doi.org/https://doi.org/10.1007/s12652-022-03772-6