Bayesian hypernetwork and Bayesian Superhypernetwork using PowerSet and nth-Powerset

Takaaki Fujita [1] *

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

Abstract

Graph theory offers a powerful framework for modeling relationships among entities; in dentistry, for example, it can represent connections between teeth or other oral structures. A *hypergraph* extends the classical graph by allowing *hyperedges* to join more than two vertices, thus capturing complex multiway interactions. A *superhypergraph* builds on this idea by introducing recursively nested powerset layers, enabling hierarchical and self-referential relationships among hyperedges. In parallel, *hypernetworks* and *superhypernetworks* generalize network models to these richer connectivity patterns. And a Bayesian network is a directed graph where nodes represent random variables and edges encode conditional dependencies via probability distributions.

In this work, we introduce the concepts of *Bayesian hypernetworks* and *Bayesian superhypernetworks*, which extend Bayesian networks by leveraging hypernetwork and superhypernetwork structures. These novel frameworks enhance the ability to model hierarchical and intricate real-world phenomena, offering significant advantages for complex decision-making and inference. We anticipate that their integration into Bayesian network theory and artificial intelligence will open new avenues for advanced probabilistic modeling and analysis.

Keywords: Bayesian hypernetwork, Bayesian Superhypernetwork, HyperGraph, Superhypergraph

1 Preliminaries

This section provides an overview of the fundamental concepts and definitions essential for the discussions in this paper. All concepts considered herein are assumed to be finite. Furthermore, unless otherwise specified, n is assumed to be a natural number.

1.1 SuperHyperGraph

In classical graph theory, a hypergraph extends the idea of a conventional graph by permitting edges—called hyperedges—to join more than two vertices. This broader framework enables the modeling of more intricate relationships between elements, thereby enhancing its utility in various fields [1, 2]. A *SuperHyperGraph* is an advanced extension of the hypergraph concept, integrating recursive powerset structures into the classical model. This concept has been recently introduced and extensively studied in the literature [3–9]. Including related concepts, we describe them below.

Definition 1.1 (Base Set)**.** A *base set* S is the foundational set from which complex structures such as powersets and hyperstructures are derived. It is formally defined as:

$$S = \{x \mid x \text{ is an element within a specified domain}\}.$$

All elements in constructs like $\mathcal{P}(S)$ or $\mathcal{P}_n(S)$ originate from the elements of S.

Definition 1.2 (Powerset)**.** The *powerset* of a set S, denoted $\mathcal{P}(S)$, is the collection of all possible subsets of S, including both the empty set and S itself. Formally, it is expressed as:

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

Definition 1.3 (n-th Powerset)**.** (cf. [10–12])

The n-th powerset of a set H, denoted $P_n(H)$, is defined iteratively, starting with the standard powerset. The recursive construction is given by:

$$P_1(H) = P(H), \quad P_{n+1}(H) = P(P_n(H)), \quad \text{for } n \geq 1.$$

Similarly, the n-th non-empty powerset, denoted $P_n^*(H)$, is defined recursively as:

$$P_1^*(H) = P^*(H), \quad P_{n+1}^*(H) = P^*(P_n^*(H)).$$

Here, $P^*(H)$ represents the powerset of H with the empty set removed.

Example 1.4 (Second Powerset and Sigma-Algebras in a Coin Toss)**.** Let the sample space for a single fair coin toss be

$$\Omega = \{\mathrm{H},\ \mathrm{T}\}.$$

The first powerset

$$P_1(\Omega) = \{\emptyset,\ \{\mathrm{H}\},\ \{\mathrm{T}\},\ \{\mathrm{H},\mathrm{T}\}\}$$

is the collection of all possible events. The second powerset

$$P_2(\Omega) = P\big(P_1(\Omega)\big)$$

is the set of all collections of events, of which the valid sigma-algebras are particular examples. For instance:

$$\mathcal{F}_{\text{trivial}} = \big\{\emptyset,\ \{\mathrm{H},\mathrm{T}\}\big\}\ \in\ P_2(\Omega),$$

$$\mathcal{F}_{\text{full}} = P_1(\Omega)\ \in\ P_2(\Omega).$$

Here $\mathcal{F}_{\text{trivial}}$ and $\mathcal{F}_{\text{full}}$ represent the trivial and full sigma-algebras on Ω. Thus, the second powerset naturally encodes the possible choices of measurable event collections in probability theory.

Definition 1.5 (Hypergraph)**.** [13, 14] A *hypergraph* $H = (V(H), E(H))$ consists of:

- A nonempty set $V(H)$ of vertices.
- A set $E(H)$ of hyperedges, where each hyperedge is a nonempty subset of $V(H)$, thereby allowing connections among multiple vertices.

Unlike standard graphs, hypergraphs are well-suited to represent higher-order relationships. In this paper, we restrict ourselves to the case where both $V(H)$ and $E(H)$ are finite.

Definition 1.6 (n-SuperHyperGraph)**.** [4, 15]
Let V_0 be a finite base set of vertices. For each integer $k \geq 0$, define the iterative powerset by

$$\mathcal{P}^0(V_0) = V_0, \quad \mathcal{P}^{k+1}(V_0) = \mathcal{P}(\mathcal{P}^k(V_0)),$$

where $\mathcal{P}(\cdot)$ denotes the usual powerset operation. An *n-SuperHyperGraph* is then a pair

$$\mathrm{SHT}^{(n)} = (V, E),$$

with

$$V \subseteq \mathcal{P}^n(V_0) \quad \text{and} \quad E \subseteq \mathcal{P}^n(V_0).$$

Each element of V is called an *n-supervertex* and each element of E an *n-superedge*.

Example 1.7 (2-SuperHyperGraph of Dice Events)**.** Let the base sample space be the outcomes of a fair six-sided die:

$$V_0 = \{1, 2, 3, 4, 5, 6\}.$$

We form the second iterated powerset $\mathcal{P}^2(V_0) = \mathcal{P}(\mathcal{P}(V_0))$, whose elements are nonempty subsets of subsets of V_0. Define the 2-SuperHyperGraph

$$\mathrm{SHT}^{(2)} = (V^{(2)},\ E^{(2)},\ w),$$

where:

$$V^{(2)} = \big\{v_1 = \{\{1,2\},\{3,4\}\},\ v_2 = \{\{2,3\},\{5,6\}\},\ v_3 = \{\{1,2\},\{5,6\}\}\big\},$$

each v_i is a 2-supervertex representing two simple events (e.g. "rolls in {1,2} or {3,4}"). We choose the 2-superedges

$$E^{(2)} = \{\, e_1 = \{v_1, v_2\},\ e_2 = \{v_2, v_3\}\},$$

linking those 2-supervertices that share a common simple event.

Finally, assign each superedge a weight equal to the joint probability of its simple-event components under the uniform die:

$$w(e_1) = P(\{1,2\} \cup \{3,4\}) \times P(\{2,3\} \cup \{5,6\}) = \tfrac{4}{6} \times \tfrac{4}{6} = \tfrac{4}{9},$$

$$w(e_2) = P(\{2,3\} \cup \{5,6\}) \times P(\{1,2\} \cup \{5,6\}) = \tfrac{4}{6} \times \tfrac{4}{6} = \tfrac{4}{9}.$$

Thus $\mathrm{SHT}^{(2)}$ captures both individual events (through its 2-supervertices) and their higher-order interactions (through 2-superedges), modeling hierarchical event collections in probability theory.

1.2 HyperNetwork and SuperhyperNetwork

A hypernetwork is a graph generalization where hyperedges connect any number of nodes, enabling modeling of multiway relationships beyond pairwise edges. An n-superhypernetwork uses vertices and hyperedges drawn from the n-th iterated powerset of a base node set to model nested, hierarchical groupings. The definitions of HyperNetwork and SuperhyperNetwork are presented below [16].

Definition 1.8 (Hypernetwork)**.** [16] A *hypernetwork* is an ordered triple

$$H = \big(V,\, \mathcal{E},\, w\big)$$

where

- V is a nonempty finite set of *nodes*;
- $\mathcal{E} \subseteq \mathcal{P}(V) \setminus \{\emptyset\}$ is the set of *hyperedges*, each hyperedge $e \in \mathcal{E}$ being a nonempty subset of nodes (allowing multi-node interactions);
- $w\colon \mathcal{E} \to \mathbb{R}_{\geq 0}$ is a *weight or attribute function* on hyperedges (omitted if unweighted).

A *directed hypernetwork* may be defined by replacing $\mathcal{E} \subseteq \mathcal{P}(V)$ with a set of *ordered* tuples of nodes or by equipping each $e \in \mathcal{E}$ with a head-tail partition. One can further add a *node-labeling* $\ell_V : V \to L_V$ and a *hyperedge-labeling* $\ell_{\mathcal{E}} : \mathcal{E} \to L_{\mathcal{E}}$ to record types or properties.

Definition 1.9 (n-SuperHypernetwork)**.** [16–18] Let V_0 be a finite base set of *nodes*. Define the n-th iterated powerset recursively by

$$\mathcal{P}^0(V_0) = V_0, \qquad \mathcal{P}^{k+1}(V_0) = \mathcal{P}\big(\mathcal{P}^k(V_0)\big) \quad (k \geq 0).$$

An *n-superhypernetwork* is a tuple

$$\mathcal{N}^{(n)} \;=\; \big(V,\, \mathcal{E},\, w\big)$$

where

- $V \subseteq \mathcal{P}^n(V_0)$ is a finite set of *n-supernodes*;
- $\mathcal{E} \subseteq \mathcal{P}^n(V_0)$ is a finite set of *n-superedges*, each superedge $e \in \mathcal{E}$ being a nonempty subset of V;
- $w\colon \mathcal{E} \to \mathbb{R}_{\geq 0}$ is an optional *weight function* assigning a nonnegative real weight (or confidence) to each superedge.

In other words, both vertices and hyperedges of the network are drawn from the n-th powerset of the base node set, capturing up to n levels of hierarchical grouping.

Example 1.10 (2-SuperHypernetwork of University Module Clusters)**.** Let the base set of courses be

$$V_0 = \{\text{Math, Physics, CS}\}.$$

We form the second iterated powerset $\mathcal{P}^2(V_0) = \mathcal{P}(\mathcal{P}(V_0))$, whose elements are nonempty collections of nonempty course-sets. Choose the following 2-supernodes:

$$\begin{aligned}
u_1 &= \big\{\{\text{Math, Physics}\}\big\},\\
u_2 &= \big\{\{\text{Physics, CS}\}\big\},\\
u_3 &= \big\{\{\text{Math, Physics}\}, \{\text{Physics, CS}\}\big\}.
\end{aligned}$$

Here:

- u_1 represents the "Math& Physics" module,
- u_2 represents the "Physics& CS" module,

- u_3 represents the combined cluster of those two modules.

Define the 2-superhypernetwork

$$\mathcal{N}^{(2)} = (V,\ \mathcal{E},\ w), \quad V = \{u_1, u_2, u_3\}, \quad \mathcal{E} = \{\{u_1, u_2\}\},$$

where the single superedge connects the two module supernodes u_1, u_2.

Assign a weight function w encoding student enrollment overlap:

$$w(\{u_1, u_2\}) = \left|\{\text{students enrolled in both modules}\}\right|.$$

For example, if 30 students take both "Math & Physics" and "Physics & CS," then $w(\{u_1, u_2\}) = 30$.

Thus $\mathcal{N}^{(2)}$ models hierarchical grouping of courses into modules and clusters, with hyperedges capturing the overlap in student participation across modules.

1.3 Bayesian Network

A Bayesian network is a directed graph where nodes represent random variables and edges encode conditional dependencies via probability distributions [19–22].

Definition 1.11 (Directed Acyclic Graph)**.** (cf. [23, 24]) A *directed acyclic graph* (DAG) is a pair $G = (V, E)$ where $V = \{X_1, \ldots, X_n\}$ is a finite set of nodes and $E \subseteq V \times V$ is a set of directed edges $X_i \to X_j$, such that no sequence of edges forms a cycle. The nodes represent random variables, and an edge $X_i \to X_j$ indicates that X_i directly influences X_j.

Definition 1.12 (Bayesian Network)**.** (cf. [25–31]) A *Bayesian network* over variables $V = \{X_1, \ldots, X_n\}$ is a pair $B = \langle G, \Theta \rangle$ where:

1. $G = (V, E)$ is a DAG.
2. For each i, $\mathrm{Pa}(X_i) \subseteq V \setminus \{X_i\}$ denotes the parents of X_i in G.
3. $\Theta = \{\theta_{X_i \mid \mathrm{Pa}(X_i)}\}$ is a collection of conditional probability distributions. Each

$$\theta_{X_i \mid \mathrm{Pa}(X_i)} : \mathrm{Dom}(\mathrm{Pa}(X_i)) \longrightarrow \Delta(\mathrm{Dom}(X_i))$$

   assigns to each configuration pa_i of the parents a probability mass function over the states of X_i.

Example 1.13 (Lung Cancer Diagnosis Bayesian Network)**.** Consider the following medical diagnosis scenario with five binary random variables:

$$V = \{\text{Smoker, Pollution, Cancer, XRay, Dyspnea}\},$$

each taking values in $\{\texttt{True}, \texttt{False}\}$. We define a Bayesian network $B = \langle G, \Theta \rangle$ as follows:

Structure G. The DAG over V has edges:

$$\text{Smoker} \to \text{Cancer}, \quad \text{Pollution} \to \text{Cancer}, \quad \text{Cancer} \to \text{XRay}, \quad \text{Cancer} \to \text{Dyspnea}.$$

Parameters Θ.

- Prior for Smoker:

$$P(\text{Smoker} = \texttt{True}) = 0.30, \quad P(\text{Smoker} = \texttt{False}) = 0.70.$$

- Prior for Pollution:

$$P(\text{Pollution} = \texttt{True}) = 0.90, \quad P(\text{Pollution} = \texttt{False}) = 0.10.$$

- Conditional for Cancer given its parents:

	Pollution = `False`	Pollution = `True`
Smoker = `False`	0.001	0.03
Smoker = `True`	0.05	0.10

  where each entry is $P(\text{Cancer} = \texttt{True} \mid \text{Smoker}, \text{Pollution})$.

- Conditional for XRay given Cancer:

$$P(\text{XRay} = \texttt{True} \mid \text{Cancer} = \texttt{True}) = 0.90, \quad P(\text{XRay} = \texttt{True} \mid \text{Cancer} = \texttt{False}) = 0.20.$$

- Conditional for Dyspnea given Cancer:

$$P(\text{Dyspnea} = \texttt{True} \mid \text{Cancer} = \texttt{True}) = 0.65, \quad P(\text{Dyspnea} = \texttt{True} \mid \text{Cancer} = \texttt{False}) = 0.30.$$

Joint Factorization. By the Bayesian network structure, the joint distribution factorizes as

$$P(\text{Smoker}, \text{Pollution}, \text{Cancer}, \text{XRay}, \text{Dyspnea}) = P(\text{Smoker})\, P(\text{Pollution})$$

$$P\big(\text{Cancer} \mid \text{Smoker}, \text{Pollution}\big)\, P\big(\text{XRay} \mid \text{Cancer}\big)$$

$$P\big(\text{Dyspnea} \mid \text{Cancer}\big).$$

This network captures how smoking and pollution jointly influence cancer risk, which in turn affects test results and symptom presence.

2 Results of this Paper

In this paper, we present the definitions and key properties of Bayesian-HyperNetworks and Bayesian n-SuperHyperNetworks.

2.1 Bayesian-HyperNetworks

A Bayesian HyperNetwork extends Bayesian networks with directed hyperedges representing conditional distributions over multiple child variables given multiple parent variables.

Definition 2.1 (Directed Acyclic Hypergraph)**.** (cf. [32, 33]) A *directed acyclic hypergraph* (DAHg) is a pair $H = (V, \mathcal{E})$ where:

- $V = \{X_1, \ldots, X_n\}$ is a finite set of nodes (random variables).
- $\mathcal{E} \subseteq \mathcal{P}(V) \times \mathcal{P}(V)$ is a set of *directed hyperedges*. Each hyperedge $(U \to W) \in \mathcal{E}$ has a nonempty *tail* $U \subseteq V$ and a nonempty *head* $W \subseteq V$, and no sequence of hyperedges forms a directed cycle.

Definition 2.2 (Bayesian HyperNetwork)**.** A *Bayesian HyperNetwork* is a pair $\mathrm{BHN} = \langle H, \Theta \rangle$ where:

- $H = (V, \mathcal{E})$ is a directed acyclic hypergraph.
- $\Theta = \{\theta_{W|U} : \mathrm{Dom}(U) \to \Delta(\mathrm{Dom}(W))\}_{(U \to W) \in \mathcal{E}}$ assigns to each hyperedge $(U \to W)$ a conditional distribution of the head variables W given the tail variables U.

The *global factorization* over all variables in V is

$$P(V) = \prod_{(U \to W) \in \mathcal{E}} P\big(W \mid U\big) = \prod_{(U \to W) \in \mathcal{E}} \theta_{W|U}(W, U).$$

Example 2.3 (Bayesian HyperNetwork for Medical Diagnosis)**.** Consider three binary random variables:

$$V = \{\text{Fever, Cough, Disease}\},$$

each taking values in {Yes,No}. We model their joint distribution with a Bayesian HyperNetwork BHN = $\langle H, \Theta \rangle$ defined as follows:

Hypergraph H.

$$H = (V, \mathcal{E}), \quad \mathcal{E} = \{\, (\{\text{Fever}\} \to \{\text{Fever}\}),\ (\{\text{Cough}\} \to \{\text{Cough}\}),\ (\{\text{Fever}, \text{Cough}\} \to \{\text{Disease}\})\}.$$

Here we use two trivial hyperedges to represent root marginals, and one nontrivial hyperedge {Fever, Cough} $\to$ {Disease} for the diagnosis.

Parameter set Θ.

- Marginal for Fever: $\theta_{\{\text{Fever}\}|\{\text{Fever}\}}$ with

$$P(\text{Fever} = \text{Yes}) = 0.10, \quad P(\text{Fever} = \text{No}) = 0.90.$$

- Marginal for Cough: $\theta_{\{\text{Cough}\}|\{\text{Cough}\}}$ with

$$P(\text{Cough} = \text{Yes}) = 0.15, \quad P(\text{Cough} = \text{No}) = 0.85.$$

- Conditional for Disease: $\theta_{\{\text{Disease}\}|\{\text{Fever,Cough}\}}$ given by

	Cough = No	Cough = Yes
Fever = No	0.01	0.05
Fever = Yes	0.30	0.90

  where each entry is $P(\text{Disease} = \text{Yes} \mid \text{Fever}, \text{Cough})$.

Global factorization. The joint probability factors as

$$P(\text{Fever}, \text{Cough}, \text{Disease}) = P(\text{Fever})\ P(\text{Cough})\ P\big(\text{Disease} \mid \text{Fever}, \text{Cough}\big).$$

This Bayesian HyperNetwork captures both simple marginals (via trivial hyperedges) and the complex dependence of disease on symptoms (via a directed hyperedge from {Fever, Cough} to {Disease}).

Theorem 2.4. *Every Bayesian network and every (undirected) hypernetwork can be embedded as a special case of a Bayesian HyperNetwork.*

Proof. **(a) Generalizing Bayesian networks.** If we restrict $\mathcal{E}$ to hyperedges $(\{X_i\} \to \{X_j\})$ of size two and let each $\theta_{\{X_j\}|\{X_i\}}$ coincide with the CPT $P(X_j \mid X_i)$, then BHN reduces exactly to a Bayesian network over the same DAG structure.

(b) Generalizing hypernetworks. If we take each $\theta_{W|U}$ to be the trivial distribution assigning probability one to a single state of W (i.e. no uncertainty), then the joint factorization imposes no probabilistic constraints, and BHN collapses to the underlying directed acyclic hypergraph H. Hence every hypernetwork (when oriented and acyclic) is recovered. □

2.2 Bayesian n-SuperHyperNetworks

A Bayesian n-SuperHyperNetwork generalizes hypernetworks further by using n-th level superedges to specify conditional distributions among hierarchical groups of variables.

Definition 2.5 (Bayesian n-SuperHyperNetwork)**.** A *Bayesian n-SuperHyperNetwork* is a pair $\mathrm{BnSHN} = \langle \mathrm{NHN}^{(n)}, \Theta \rangle$ where:

- $\mathrm{NHN}^{(n)} = (V, \mathcal{E}, w)$ is an n-SuperHyperNetwork.
- $\Theta = \{\theta_{W|U} : \mathrm{Dom}(U) \to \Delta(\mathrm{Dom}(W))\}_{(U\to W)\in\mathcal{E}}$ assigns to each n-superedge $(U \to W)$ a conditional distribution over the n-supernodes W given U.

The joint distribution factorizes as

$$P(V) = \prod_{(U\to W)\in\mathcal{E}} P(W \mid U).$$

Example 2.6 (Bayesian 2-SuperHyperNetwork for Symptom Cluster Diagnosis)**.** Let the base set of binary symptoms be

$$V_0 = \{\text{Fever, Cough, Fatigue}\},$$

each taking values in $\{\texttt{Present}, \texttt{Absent}\}$. The second powerset $\mathcal{P}^2(V_0)$ consists of nonempty subsets of nonempty subsets of V_0. We select the following 2-supervertices:

$$V = \{u_1 = \{\{\text{Fever}\}, \{\text{Cough}\}\}, \quad u_2 = \{\{\text{Cough}\}, \{\text{Fatigue}\}\}, \quad u_3 = \{\{\text{Fever}\}, \{\text{Fatigue}\}\}\}.$$

Each u_i represents a "two-symptom cluster" event. Define a single 2-superedge:

$$e = \{u_1, u_2\} \longrightarrow \{u_3\},$$

indicating that the joint presence of clusters $\{\text{Fever}, \text{Cough}\}$ and $\{\text{Cough}, \text{Fatigue}\}$ influences the cluster $\{\text{Fever}, \text{Fatigue}\}$.

We assign a weight function $w(e) = 1$ (ignored for inference) and specify the conditional distribution

$$\theta_{\{u_3\}|\{u_1,u_2\}} : \{\texttt{Present}, \texttt{Absent}\}^2 \longrightarrow \Delta(\{\texttt{Present}, \texttt{Absent}\})$$

as follows:

$$\begin{aligned}
P(u_3 = \texttt{Present} \mid u_1 = \texttt{Present},\, u_2 = \texttt{Present}) &= 0.80,\\
P(u_3 = \texttt{Present} \mid u_1 = \texttt{Present},\, u_2 = \texttt{Absent}) &= 0.30,\\
P(u_3 = \texttt{Present} \mid u_1 = \texttt{Absent},\, u_2 = \texttt{Present}) &= 0.25,\\
P(u_3 = \texttt{Present} \mid u_1 = \texttt{Absent},\, u_2 = \texttt{Absent}) &= 0.05,
\end{aligned}$$

with the complementary probability for $\texttt{Absent}$. The joint distribution over all 2-supervertices factorizes as

$$P(u_1, u_2, u_3) = P(u_1)\; P(u_2)\; P(u_3 \mid u_1, u_2),$$

where $P(u_1)$ and $P(u_2)$ are prior marginals (e.g. set to 0.2–0.4 based on epidemiological data).

This Bayesian 2-SuperHyperNetwork models how overlapping two-symptom clusters combine to influence the emergence of another cluster, capturing hierarchical interactions in medical diagnosis.

Example 2.7 (Bayesian 2-SuperHyperNetwork in Supply Chain Reliability)**.** Let the base set of suppliers be

$$V_0 = \{\text{S1, S2, S3}\},$$

each with a binary state $\{\texttt{Up}, \texttt{Down}\}$. We form the second iterated powerset $\mathcal{P}^2(V_0)$ and choose three 2-supervertices:

$$\begin{aligned}
u_1 &= \{\{\text{S1}\},\ \{\text{S2}\}\},\\
u_2 &= \{\{\text{S2}\},\ \{\text{S3}\}\},\\
u_3 &= \{u_1,\ u_2\}.
\end{aligned}$$

Here u_1 represents the cluster of suppliers S1 and S2, u_2 the cluster of S2 and S3, and u_3 the meta-cluster combining those two clusters.

Define the directed 2-superhypergraph

$$\mathrm{NHN}^{(2)} = \big(V, \mathcal{E}, w\big), \quad V = \{u_1, u_2, u_3\}, \quad \mathcal{E} = \{\, (\{u_1, u_2\} \to \{u_3\})\},$$

with trivial weight $w \equiv 1$.

Assign the parameter set Θ as follows:

- Prior reliability of cluster u_1: $\theta_{u_1|\emptyset}$: $P(u_1 = \mathtt{Up}) = 0.90$, $P(u_1 = \mathtt{Down}) = 0.10$.
- Prior reliability of cluster u_2: $\theta_{u_2|\emptyset}$: $P(u_2 = \mathtt{Up}) = 0.85$, $P(u_2 = \mathtt{Down}) = 0.15$.
- Conditional reliability of meta-cluster u_3 given u_1, u_2:

  $$\theta_{u_3|\{u_1,u_2\}}(u_3,\, u_1, u_2) = P\big(u_3 = \mathtt{Up} \mid u_1, u_2\big),$$

  with

	$u_2 = \mathtt{Down}$	$u_2 = \mathtt{Up}$
$u_1 = \mathtt{Down}$	0.20	0.50
$u_1 = \mathtt{Up}$	0.60	0.95

  and complementary probabilities for $u_3 = \mathtt{Down}$.

The joint distribution factorizes as

$$P(u_1, u_2, u_3) = P(u_1)\; P(u_2)\; P\big(u_3 \mid u_1, u_2\big),$$

which models how individual cluster reliabilities propagate through the hierarchical network to determine overall system reliability.

Example 2.8 (Bayesian 2-SuperHyperNetwork for Financial Portfolio Risk)**.** Let the base set of assets be

$$V_0 = \{\mathrm{A},\ \mathrm{B},\ \mathrm{C}\},$$

each with binary return state $\{\mathtt{High}, \mathtt{Low}\}$. We form the second iterated powerset $\mathcal{P}^2(V_0)$ and select three 2-supervertices:

$$u_1 = \big\{\{\mathrm{A}\}, \{\mathrm{B}\}\big\},$$
$$u_2 = \big\{\{\mathrm{B}\}, \{\mathrm{C}\}\big\},$$
$$u_3 = \big\{\{\mathrm{A}\}, \{\mathrm{C}\}\big\}.$$

Here u_1 represents the cluster of assets A and B, u_2 the cluster of B and C, and u_3 the cluster of A and C.

Define the directed 2-superhypergraph

$$\mathrm{NHN}^{(2)} = \big(V, \mathcal{E}, w\big), \quad V = \{u_1, u_2, u_3\}, \quad \mathcal{E} = \{\, (\{u_1, u_2\} \to \{u_3\})\},$$

with uniform weight $w \equiv 1$.

Assign the parameter set Θ as follows:

- Prior marginal for u_1:
  $$P\big(u_1 = \mathtt{High}\big) = 0.55, \quad P\big(u_1 = \mathtt{Low}\big) = 0.45.$$
- Prior marginal for u_2:
  $$P\big(u_2 = \mathtt{High}\big) = 0.60, \quad P\big(u_2 = \mathtt{Low}\big) = 0.40.$$

- Conditional distribution for u_3 given u_1, u_2:

$$\theta_{\{u_3\}|\{u_1,u_2\}} = P\big(u_3 = \texttt{High} \mid u_1, u_2\big),$$

specified by

	$u_2 = \texttt{Low}$	$u_2 = \texttt{High}$
$u_1 = \texttt{Low}$	0.20	0.50
$u_1 = \texttt{High}$	0.65	0.95

with complementary probabilities for $u_3 = \texttt{Low}$.

The joint probability over all 2-supervertices factorizes as

$$P(u_1, u_2, u_3) = P(u_1)\; P(u_2)\; P\big(u_3 \mid u_1, u_2\big).$$

This Bayesian 2-SuperHyperNetwork models how the performance clusters of assets A/B and B/C combine to influence the risk cluster of A/C, capturing hierarchical dependencies in portfolio risk assessment.

Theorem 2.9. *Bayesian n-SuperHyperNetworks simultaneously generalize Bayesian networks, Bayesian HyperNetworks, and n-SuperHyperNetworks.*

Proof. BnSHN reduces to:

- A Bayesian HyperNetwork when $n = 1$ (since $\P^1(V_0) = \P(V_0)$).
- A Bayesian network when $n = 1$ and all hyperedges have tails and heads of size one.
- An n-SuperHyperNetwork when each $\theta_{W|U}$ is trivial (unit mass).

Thus all three structures arise as special cases. □

Theorem 2.10 (Local Markov Property)**.** *Let* BnSHN $= \langle \mathrm{NHN}^{(n)}, \Theta \rangle$ *be a Bayesian n-SuperHyperNetwork on supernodes V and superedges $\mathcal{E}$. For any supernode $W \in V$ with parent supernodes* $\mathrm{Pa}(W) = U$*, define the set of non-descendant, non-parent supernodes as*

$$\mathrm{ND}(W) \;=\; V \;\setminus\; \big(\mathrm{Desc}(W) \,\cup\, \mathrm{Pa}(W)\big).$$

Then

$$W \;\perp\!\!\!\perp\; \mathrm{ND}(W) \;\Big|\; \mathrm{Pa}(W)$$

in the joint distribution $P(V) = \prod_{(U \to X) \in \mathcal{E}} P(X \mid U)$.

Proof. By construction, the joint density factorizes as

$$P(V) = \prod_{(U \to X) \in \mathcal{E}} P\big(X \mid U\big).$$

Consider the conditional distribution of W given all other supernodes:

$$P\big(W \mid V \setminus \{W\}\big) = \frac{P(V)}{\sum_w P(V)} = P\big(W \mid \mathrm{Pa}(W)\big),$$

because only the factor $P(W \mid \mathrm{Pa}(W))$ involves W; all other factors either do not depend on W or cancel in the normalization. Hence, once $\mathrm{Pa}(W)$ is known, W is independent of every other supernode not in $\mathrm{Desc}(W)$ or $\mathrm{Pa}(W)$, i.e. $W \perp\!\!\!\perp \mathrm{ND}(W) \mid \mathrm{Pa}(W)$. □

Theorem 2.11 (Global Markov Property)**.** *Let $A, B, C \subseteq V$ be disjoint sets of supernodes. Form the* moral graph *of* $\mathrm{NHN}^{(n)}$ *by*

1. *Replacing each directed hyperedge* $(U \to W)$ *by undirected edges between every pair in* $U \cup W$.
2. *"Marrying" all parents in U (making U a clique) and dropping edge directions.*

If every undirected path from any $a \in A$ *to any* $b \in B$ *intersects* C, *then in the joint distribution,*

$$A \perp\!\!\!\perp B \,\big|\, C.$$

Proof. The moral graph construction ensures that all conditional dependencies implied by directed hyperedges are represented as undirected connections. A classic separation argument in undirected graphs then applies: if C blocks all paths between A and B, no factor in the factorization can transmit probabilistic influence between them once C is observed. Therefore $P(A, B \mid C) = P(A \mid C)\, P(B \mid C)$, establishing $A \perp\!\!\!\perp B \mid C$. □

Theorem 2.12 (Parameter Independence under Dirichlet Priors)**.** *Assume each conditional distribution* $\theta_{W|U}$ *in* Θ *has a conjugate Dirichlet prior*

$$\theta_{W|U} \ \sim \ \mathrm{Dir}\big(\alpha_{w|u} : w \in \mathrm{Dom}(W)\big).$$

Given independent data counts $N_{w,u}$ *for each configuration* $(U = w)$, *the posterior factorizes as*

$$p(\Theta \mid \{N_{w,u}\}) = \prod_{(U \to W) \in \mathcal{E}} \mathrm{Dir}\big(\alpha_{w|u} + N_{w,u} : w \in \mathrm{Dom}(W)\big).$$

Proof. Because the joint likelihood factorizes $\prod_{(U \to W)} \prod_{w,u} \theta_{W|U}(w \mid u)^{N_{w,u}}$, and each prior over $\theta_{W|U}$ is Dirichlet, posterior conjugacy implies

$$p(\theta_{W|U} \mid \{N_{w,u}\}) \ \propto \ \Big[\prod_{w,u} \theta_{W|U}(w \mid u)^{\alpha_{w|u}-1}\Big]\Big[\prod_{w,u} \theta_{W|U}(w \mid u)^{N_{w,u}}\Big] = \mathrm{Dir}\big(\alpha_{w|u} + N_{w,u}\big).$$

Since factors share no parameters, the full posterior factorizes across hyperedges. □

Theorem 2.13 (Consistency under Aggregation)**.** *If two Bayesian n-SuperHyperNetworks* BnSHN_1 *and* BnSHN_2 *share the same hypergraph structure but differ in variable granularity—where* BnSHN_2 *coarsens supernodes of* BnSHN_1*—then marginalizing the finer network over the aggregated states recovers the coarser network's joint distribution.*

Proof. Aggregation corresponds to summing out detailed supernodes in the factorized distribution of BnSHN_1. Because the factorization aligns with the coarser network's hyperedges (which are unions of fine-level hyperedges), the resulting marginalized distribution has the same factorization as BnSHN_2, ensuring consistency. □

Funding

This study did not receive any financial or external support from organizations or individuals.

Acknowledgments

We extend our sincere gratitude to everyone who provided insights, inspiration, and assistance throughout this research. We particularly thank our readers for their interest and acknowledge the authors of the cited works for laying the foundation that made our study possible. We also appreciate the support from individuals and institutions that provided the resources and infrastructure needed to produce and share this paper. Finally, we are grateful to all those who supported us in various ways during this project.

Data Availability

This research is purely theoretical, involving no data collection or analysis. We encourage future researchers to pursue empirical investigations to further develop and validate the concepts introduced here.

Ethical Approval

As this research is entirely theoretical in nature and does not involve human participants or animal subjects, no ethical approval is required.

Conflicts of Interest

The authors confirm that there are no conflicts of interest related to the research or its publication.

Disclaimer

This work presents theoretical concepts that have not yet undergone practical testing or validation. Future researchers are encouraged to apply and assess these ideas in empirical contexts. While every effort has been made to ensure accuracy and appropriate referencing, unintentional errors or omissions may still exist. Readers are advised to verify referenced materials on their own. The views and conclusions expressed here are the authors' own and do not necessarily reflect those of their affiliated organizations.

References

[1] Derun Cai, Moxian Song, Chenxi Sun, Baofeng Zhang, Shenda Hong, and Hongyan Li. Hypergraph structure learning for hypergraph neural networks. In *IJCAI*, pages 1923–1929, 2022.

[2] Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao. Hypergraph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 3558–3565, 2019.

[3] Takaaki Fujita and Florentin Smarandache. Fundamental computational problems and algorithms for superhypergraphs. *HyperSoft Set Methods in Engineering*, 3:32–61, 2025.

[4] Florentin Smarandache. *Extension of HyperGraph to n-SuperHyperGraph and to Plithogenic n-SuperHyperGraph, and Extension of HyperAlgebra to n-ary (Classical-/Neutro-/Anti-) HyperAlgebra*. Infinite Study, 2020.

[5] Y. V. M. Cepeda, M. A. R. Guevara, E. J. J. Mogro, and R. P. Tizano. Impact of irrigation water technification on seven directories of the san juan-patoa river using plithogenic n-superhypergraphs based on environmental indicators in the canton of pujili, 2021. *Neutrosophic Sets and Systems*, 74:46–56, 2024.

[6] Berrocal Villegas Salomón Marcos, Montalvo Fritas Willner, Berrocal Villegas Carmen Rosa, Flores Fuentes Rivera María Yissel, Espejo Rivera Roberto, Laura Daysi Bautista Puma, and Dante Manuel Macazana Fernández. Using plithogenic n-superhypergraphs to assess the degree of relationship between information skills and digital competencies. *Neutrosophic Sets and Systems*, 84:513–524, 2025.

[7] Takaaki Fujita and Florentin Smarandache. A concise study of some superhypergraph classes. *Neutrosophic Sets and Systems*, 77:548–593, 2024.

[8] Eduardo Martín Campoverde Valencia, Jessica Paola Chuisaca Vásquez, and Francisco Ángel Becerra Lois. Multineutrosophic analysis of the relationship between survival and business growth in the manufacturing sector of azuay province, 2020–2023, using plithogenic n-superhypergraphs. *Neutrosophic Sets and Systems*, 84:341–355, 2025.

[9] Florentin Smarandache. *Introduction to the n-SuperHyperGraph-the most general form of graph today*. Infinite Study, 2022.

[10] Florentin Smarandache. Foundation of superhyperstructure & neutrosophic superhyperstructure. *Neutrosophic Sets and Systems*, 63(1):21, 2024.

[11] F. Smarandache. Introduction to superhyperalgebra and neutrosophic superhyperalgebra. *Journal of Algebraic Hyperstructures and Logical Algebras*, 2022.

[12] Florentin Smarandache. *SuperHyperFunction, SuperHyperStructure, Neutrosophic SuperHyperFunction and Neutrosophic SuperHyperStructure: Current understanding and future directions*. Infinite Study, 2023.

[13] Alain Bretto. Hypergraph theory. *An introduction. Mathematical Engineering. Cham: Springer*, 1, 2013.

[14] Claude Berge. *Hypergraphs: combinatorics of finite sets*, volume 45. Elsevier, 1984.

[15] Florentin Smarandache. n-superhypergraph and plithogenic n-superhypergraph. *Nidus Idearum*, 7:107–113, 2019.

[16] Takaaki Fujita. Exploration of graph classes and concepts for superhypergraphs and n-th power mathematical structures. *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond*, 3(4):512, 2025.

[17] Takaaki Fujita. Hypergraph and superhypergraph approaches in electronics: A hierarchical framework for modeling power-grid hypernetworks and superhypernetworks. *Journal of Energy Research and Reviews*, 17(6):102–136, 2025.

[18] Shouxian Zhu. Neutrosophic n-superhypernetwork: A new approach for evaluating short video communication effectiveness in media convergence. *Neutrosophic Sets and Systems*, 85:1004–1017, 2025.

[19] Olivier Pourret, Patrick Na, Bruce Marcot, et al. *Bayesian networks: a practical guide to applications*. John Wiley & Sons, 2008.

[20] Radhakrishnan Nagarajan, Marco Scutari, and Sophie Lèbre. Bayesian networks in r. *Springer*, 122:125–127, 2013.

[21] Zoubin Ghahramani. Learning dynamic bayesian networks. *International School on Neural Networks, Initiated by IIASS and EMFCSC*, pages 168–197, 1997.

[22] David Heckerman. Bayesian networks for data mining. *Data mining and knowledge discovery*, 1:79–119, 1997.

[23] Peter WG Tennant, Eleanor J Murray, Kellyn F Arnold, Laurie Berrie, Matthew P Fox, Sarah C Gadd, Wendy J Harrison, Claire Keeble, Lynsie R Ranker, Johannes Textor, et al. Use of directed acyclic graphs (dags) to identify confounders in applied health research: review and recommendations. *International journal of epidemiology*, 50(2):620–632, 2021.

[24] Jean C Digitale, Jeffrey N Martin, and Medellena Maria Glymour. Tutorial on directed acyclic graphs. *Journal of Clinical Epidemiology*, 142:264–267, 2022.

[25] Yuansen Wang, Guibin Lv, Jialin He, Feng Cheng, and Dongke Li. Dynamic evaluation of road and bridge engineering construction safety risk based on fuzzy dynamic bayesian network research. *Journal of Intelligent & Fuzzy Systems*, (Preprint):1–12, 2024.

[26] Yanfeng Li, Hongzhong Huang, Jinhua Mi, Weiwen Peng, and X. M. Han. Reliability analysis of multi-state systems with common cause failures based on bayesian network and fuzzy probability. *Annals of Operations Research*, 311:195 – 209, 2019.

[27] Qi Zhang, Chunjie Zhou, Yu-Chu Tian, Naixue N. Xiong, Yuanqing Qin, and Bowen Hu. A fuzzy probability bayesian network approach for dynamic cybersecurity risk assessment in industrial control systems. *IEEE Transactions on Industrial Informatics*, 14:2497–2506, 2018.

[28] Yong Bo Xuan, Chang Qiang Huang, and Wang Xi Li. Air combat situation assessment by gray fuzzy bayesian network. *Applied Mechanics and Materials*, 69:114–119, 2011.

[29] Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine learning*, 29:131–163, 1997.

[30] Todd Andrew Stephenson. An introduction to bayesian network theory and usage. 2000.

[31] David Heckerman. A tutorial on learning with bayesian networks. *Learning in graphical models*, pages 301–354, 1998.

[32] Mohammad Ali Javidian, Zhiyu Wang, Linyuan Lu, and Marco Valtorta. On a hypergraph probabilistic graphical model. *Annals of Mathematics and Artificial Intelligence*, 88:1003–1033, 2020.

[33] Merten Popp, Sebastian Schlag, Christian Schulz, and Daniel Seemaier. Multilevel acyclic hypergraph partitioning. In *2021 Proceedings of the Workshop on Algorithm Engineering and Experiments (ALENEX)*, pages 1–15. SIAM, 2021.