Boolean Ring of Blackening Operators and applications to Graph Learning

Billal Chouli*, Samuel Guilluy †, and Florian Méhats ‡

Abstract. This paper presents a generalization of Hammersley and Clifford's random field theorem by using the theory of Boolean rings and principal ideals. Our proposed methodology allows for the construction of graphical models that can efficiently learn other data structures, such as trees, with fewer assumptions and still maintain computational tractability. This new approach relaxes the Markovian assumption of traditional random field models, providing a more flexible and powerful tool for graph representation learning.

Key words. Boolean Ring, Markov Property, Propositional Logic

MSC codes. 60J20, 03B70, 03G05

1. Introduction.

The field of graph structured data is increasingly important in a variety of fields, including computational biology (27), natural language processing (14), and social science (26). One of the main challenges in this area is understanding the underlying structure of the data and the analysis of the functions defined on these data, as highlighted by (2) in their 2017 paper on geometric deep learning.

Markov Random Fields, first introduced by Hammersley and Clifford in 1971(6), have been a fundamental tool for understanding data structures, providing factorization formulae for Markovian lattices. However, these models have limitations in terms of expressiveness, due to the strong assumption on dependencies between neighbours, and recent studies (25) (23) have shown the limitations of Graph Neural Networks (GNNs) as well.

The objective of this paper is to propose an algebraic approach to the problem of finding the best neural network architecture given the underlying data structure and dependency relationships involved. We provide a new methodology for exploiting data structure and relational dependency, which relaxes the Markovian assumption of Conditional Random Fields (CRFs).

We base our analysis on an extension of Hammersley and Clifford's theorem (6) using the theory of Boolean rings presented by (20). Specifically, we exploit the property of principal ideals in a Boolean ring to find the best factorization of the join probability function that respects certain types of dependency constraints. This approach allows us to overcome the limitations of existing models and has the potential to greatly impact a wide range of fields that rely on graph structured data.

In this paper, we present a detailed analysis of the use of Boolean rings of blackening operators as a tool for graph representation learning. We begin by providing a comprehensive overview of the theory of Boolean rings, highlighting its connection to Boolean algebra and the concept of principal ideals. We then introduce the concept of blackening operators and demonstrate how they can be used to factorize the join probability function in a graph.

*Headmind Partners AI & Blockchain, https://www.headmind.com/fr/.

†Université Rennes 1, IRMAR.

‡Ravel Technologies, on leave from University of Rennes 1, IRMAR.

Building on this, we propose a new demonstration of the Hammersley and Clifford's theorem, which is based on the theory of Boolean rings. This approach serves as a novel methodology for exploiting data structure and relational dependency in graph representation learning. To further showcase the effectiveness of our proposed approach, we provide an example of its application in the context of arborescences.

2. Related works.

Markov Random Fields. Markov Random Fields (MRFs) have been a fundamental tool for understanding data structures since their introduction by Hammersley and Clifford (6) in 1971. They provide a way to factorize graphs under the Markovian assumption, and have been successfully applied in various fields such as Natural Language Processing (12) (8) and Computer Vision (3). However, MRFs have limitations in terms of expressiveness due to the strong assumption of dependencies between neighbours, which has been highlighted in previous research studies (4) (17) (5).

Graph Neural Networks. Graph Neural Networks (GNNs) have emerged as a powerful tool for various tasks related to graph structured data. They offer a wide range of dependencies to be learned through the use of attention mechanisms (GAT) (24) or convolutional mechanisms (GCN) (11), and have the advantage of allowing for better parallelization of computation between different GPUs. However, recent studies (23) have shown that GNNs have limitations in terms of expressiveness and may have negative effects such as under-reaching, over-smoothing and over-squashing. In particular, they are at most as expressive as the Weisfeiler-Lehman test for graphs isomorphisms (25).

Boolean algebra of projectors. Boolean algebra of projectors, as introduced by Stone (20) in his representation theorem for Boolean algebras, provides a powerful tool for understanding data structures. Stone's theorem states that every Boolean algebra is isomorphic to a certain field of sets, and this result has been widely used in various fields such as the spectral theory of operators on a Hilbert space (1), Boolean algebra of projectors (10) (16) and lattices structures (15) (18). These results also lead to many results in category theory related to topological space (7).

Tree structure and oriented graph. Working with oriented graphs poses several challenges, particularly in terms of the representation and analysis of their underlying structure. One of the main challenges is the difficulty to interpret the spectral domain construction of non-symmetric Laplacian matrices, which often leads to treating directed graphs as undirected and losing important information. Many efforts have been made to address these challenges, including the use of Graph Neural Networks (GNNs) with architectures such as BiGraphSAGE (13), LSTM (21), and Neural Trees (22). These architectures aim to exploit the tree structure present in directed graphs, allowing for more accurate and efficient representations of the graph.

3. Preliminaries: Boolean rings and ideal factorization. The main goal of this section is to provide an algebraic framework for the rest of the paper. More precisely, we define conditions under which unions and intersections of certain ideals can be reduced into more compact forms. An abstract formulation of the $I(\beta)$ used in (6) for the CRF demonstration is stated and proved in Theorem 3.9 using elementary results from the theory of Boolean rings and principal ideals. This theorem will then be applied in Section 4.

Let us start with a few useful definitions.

Definition 3.1. *Let A be a commutative ring for the operations $\oplus$ and $\otimes$. We say that A is a* Boolean ring *if all the elements of A are idempotent, i.e.*

$$\forall a \in A, \quad a \otimes a = a.$$

On the Boolean ring A, we shall consider the partial order $\leq$ defined by

$$\forall (b,c) \in A^2, \qquad b \leq c \quad \text{if, and only if,} \quad b \otimes c = b. \tag{3.1}$$

In the sequel, 0 and $\mathbb{1}$ respectively denote the neutral elements for $\oplus$ and $\otimes$.

Definition 3.2. *A non empty subset I of a Boolean ring is an* ideal *if, and only if,*

- *I is closed under the addition:*

$$\forall (a,b) \in I^2, \quad a \oplus b \in I;$$

- *I is stable with respect to the partial order $\leq$ in the following sense:*

$$\forall a \in I, \quad \forall c \in A, \quad \text{if} \quad c \leq a \quad \text{then} \quad c \in I.$$

Definition 3.3. *An ideal α of a Boolean Ring A is said to be* principal *if it is generated by one of its elements:*

$$\exists a \in \alpha \quad \text{such that} \quad \alpha = I(a) := \{c \in A : c \leq a\}.$$

Definition 3.4. *Two elements a and b of a Boolean ring A are said to be* orthogonal *if $a \otimes b = 0$ and two non empty subsets α and β of A are said to be* orthogonal *if every element of α is orthogonal to every element of β. We also denote it by $\alpha \perp \beta$. For all subset α of A, we denote*

$$\alpha^{\perp} = \{b \in A : \forall a \in \alpha, \, a \otimes b = 0\}.$$

Lemma 3.5. *For all a in A, there exists a unique b in A such that*

$$a \otimes b = 0 \quad \text{and} \quad a \oplus b = \mathbb{1}.$$

This element, denoted by a', is called the complementary *of a.*

Let us now introduce the notion of Boolean algebra, a mathematical structure that is isomorphic to Boolean rings but defined using the meet $\wedge$ and join $\vee$ operators instead of the $\oplus$ and $\otimes$ operators. We use the definition from (16) based on partially orders sets and lattices, which is equivalent to the definition based on $\vee$ and $\wedge$ presented in (9) and used by (19).

To begin with, let us recall a few classical definitions. Let $(\mathcal{O}, \leq)$ be a partially ordered set and $J \subset \mathcal{O}$. An element $a \in \mathcal{O}$ is an *upper bound* (resp. a *lower bound*) of J if, for all $b \in J$, we have $b \leq a$ (resp. $a \leq b$). An upper bound (resp. a lower bound) $a \in \mathcal{O}$ of J is said to be a *least upper bound* (resp. a *greatest lower bound*) of J if every upper bound (resp. *lower bound*) c of J satisfies $a \leq c$ (resp. $c \leq a$).

Definition 3.6. *A partially order set* $\mathbb{L}$ *is called a* lattice *if every pair* $(x, y) \in \mathbb{L}^2$ *has a least upper bound and a greatest lower bound, respectively denoted by* $x \vee y$ *and* $x \wedge y$*. Moreover,*

1. *the lattice* $\mathbb{L}$ *is said to be* distributive *if, for all* $(x, y, z) \in \mathbb{L}^3$*, we have* $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ *and* $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$*;*
2. *the lattice is said to have a unit if there exists a unique* $\mathbb{1} \in \mathbb{L}$ *such that* $x \leq \mathbb{1}$ *for all* $x \in L$*;*
3. *the lattice is said to have a zero if there exists a unique element* $0 \in \mathbb{L}$ *such that* $0 \leq x$ *for all* $x \in \mathbb{L}$*;*
4. *the lattice is complemented if it has a unit and a zero and if, for every* $x \in \mathbb{L}$*, there is an element* $x' \in \mathbb{L}$ *(called the complement of* x*) such that* $x \wedge x' = 0$ *and* $x \vee x' = \mathbb{1}$*.*

Definition 3.7. *A distributive and complemented lattice is called a* Boolean algebra*.*

With the notations of Definition 3.6, we can identify a Boolean algebra by a quintuple:

$$(\mathbb{L}, \vee, \wedge, 0, \mathbb{1}).$$

The following proposition (whose proof can be found in (19)) makes a link between Boolean rings and Boolean algebras.

Proposition 3.8. *Let* B *be a Boolean algebra equipped with the operators* $\vee$ *and* $\wedge$*. Then* B *can be converted into a Boolean ring with respect to the addition* $\oplus$ *and the multiplication* $\otimes$ *defined by*

$$a \oplus b = (a \wedge b') \vee (a' \wedge b) \quad \text{and} \quad a \otimes b = a \wedge b.$$

Conversely, a Boolean ring with partial order defined by Equation (3.1) is a Boolean algebra and we have

$$a \vee b = a \oplus b \oplus a \otimes b \quad \text{and} \quad a \wedge b = a \otimes b.$$

Hence, one can consider the three operations $\oplus$, $\wedge$ (equivalent to $\otimes$) and $\vee$ as three operations operating on a Boolean ring. We also define the addition of subsets of a Boolean Ring as

$$I + J = \{i \oplus j \mid i \in I \text{ and } j \in J\}.$$

We can now state and prove the main theorem of this section (together with Corollary 3.11).

Theorem 3.9. *Let* $(a_j)_{j\in J}$ *and* $(b_j)_{j\in J}$ *be two sets of elements of a Boolean ring* A*. Then we have*

$$I\left(\prod_{j\in J}(a_j \vee b_j)\right) = \sum_{K\subset\mathcal{P}(J)} \bigcap_{k_1\in K} I(a_{k_1}) \bigcap_{k_2\in J\setminus K} I(b_{k_2}) \bigcap_{k_3\in J\setminus K} I(a'_{k_3}),$$

where $\mathcal{P}(J)$ *is the set of all the subsets of* J *and* $\sum$ *is the addition on sets.*

Before proving this theorem, we recall Theorem 31 of (20).

Lemma 3.10. *The class* $\mathbb{P}$ *of all principal ideals in a Boolean ring* A *is isomorphic to the Boolean ring* A *itself in accordance with the following relations :*

1. $I(a) = I(b)$ *if and only if* $a = b$*.*
2. $I(a \oplus b) = I(a) + I(b) = (I(a)I(b)^{\perp}) \cup I(a)^{\perp}I(b)$*.*
3. $I(a \vee b) = I(a) \cup I(b)$*.*

4. $I(a \wedge b) = I(a) \cap I(b)$.
5. $I(a') = I(a)^{\perp}$.

Proof of Theorem 3.9. We first deduce from Lemma 3.10 that

$$\begin{aligned} I(a \vee b) &= I(a \oplus b \oplus a \otimes b), \\ &= I(a \oplus b \otimes (\mathbb{1} \oplus a)) \\ &= I(a \oplus b \otimes a'), \\ &= I(a) + I(b \otimes a'), \\ &= I(a) + I(b) \cap I(a'), \\ &= I(a) + I(b) \cap I(a)^{\perp}. \end{aligned} \tag{3.2}$$

We can now develop, using again Lemma 3.10,

$$\begin{aligned} I\left(\prod_{j \in J}(a_j \vee b_j)\right) &= \bigcap_{j \in J}(I(a_j) + I(b_j) \cap I(a_j)^{\perp}), \\ &= \sum_{K \subset P(J)} \bigcap_{k_1 \in K} I(a_{k_1}) \bigcap_{k_2 \in J \setminus K} I(b_{k_2}) \cap I(a_{k_2})^{\perp}, \\ &= \sum_{K \subset P(J)} \bigcap_{k_1 \in K} I(a_{k_1}) \bigcap_{k_2 \in J \setminus K} I(b_{k_2}) \bigcap_{k_3 \in J \setminus K} I(a'_{k_3}), \end{aligned} \tag{3.3}$$

which yields the result. ∎

We now study the conditions under which each term of the sum does not vanish.

Corollary 3.11. *With the same notations as in Theorem 3.9, consider a subset $K \subset J$. Then the term $\left(\bigcap_{k_1 \in K} I(a_{k_1}) \bigcap_{k_2 \in J \setminus K} I(b_{k_2}) \bigcap_{k_3 \in J \setminus K} I(a'_{k_3})\right)$ is equal to $\{0\}$ if one of the six following conditions are not satisfied:*

1. $\forall (k_1, k_1') \in K^2, \quad a_{k_1} a_{k_1'} \neq 0$,
2. $\forall (k_2, k_2') \in (J \setminus K)^2, \quad b_{k_2} b_{k_2'} \neq 0$,
3. $\forall (k_3, k_3') \in (J \setminus K)^2, \quad a'_{k_3} a'_{k_3'} \neq 0$,
4. $\forall k_1 \in K, \forall k_2 \in J \setminus K, \quad a_{k_1} b_{k_2} \neq 0$,
5. $\forall k_1 \in K, \forall k_3 \in J \setminus K, \quad a_{k_1} \nleq a_{k_3}$,
6. $\forall (k_2, k_3) \in J \setminus K, \quad b_{k_2} \nleq a_{k_3}$.

The proof of this corollary is in Appendix A. This general factorization formula will enable to study some specific cases of a_k and b_k in Section 5. However, to apply this result, we first need to introduce a framework in relation with graphs and compatible with Boolean rings.

4. The Boolean ring of blackening operators.

The main goal of this section is the construction of the Boolean ring of blackening projectors. We will first present the structure and then demonstrate that this is a Boolean ring.

Definition 4.1. *Let G be a (oriented or not) graph with $Z = \{z_i\}$ the set of nodes of G and $E = (z_i, z_j)$ the set of the edges of G. We define the* colors *C as a finite set of elements $\{c_j\}$*

containing the color "black". A coloration *of a graph G is an application χ from Z to C.*

$$\begin{aligned}\chi\colon Z &\to C\\ z_i &\mapsto \chi(z_i) = c_j\end{aligned}$$

We also introduce the notation χ_Y to be the application that attributes the same color to the graph as χ except for the set of nodes Y which are blacken. In particular, χ_Z corresponds to a totally black coloration of the graph. The set of colorations will be denoted by $\mathcal{C}$.

We now consider the set $\mathcal{F}$ of real-valued functions defined on the colorations $\mathcal{C}$ of the graph G and which attribute the value zero to χ_Z. Our main quantity of interest will be the set of operators on $\mathcal{F}$, on which we first define three operations $\vee$, $\wedge$ and $'$. We denote by $\mathbb{1}$ and 0 respectively the identity operator and the null operator.

Definition 4.2. *Let P and Q be two operators on $\mathcal{F}$, we define:*

$$\begin{aligned}P \vee Q &= P + Q - P \circ Q,\\ P \wedge Q &= P \circ Q,\\ P' &= \mathbb{1} - P,\end{aligned} \tag{4.1}$$

where $+$ and $-$ are induced by the corresponding operations on $\mathbb{R}$ and $\circ$ is the composition operator.

Now we consider a specific class of operators on $\mathcal{F}$ (illustrated below on Figure 1) and the ring generated by this class.

Definition 4.3. *Considering a subset Y of Z, we define the operator B_Y, called the* pure blackening operator*, as the following operator on $\mathcal{F}$:*

$$\forall F \in \mathcal{F},\, \forall \chi \in \mathcal{C}, \qquad (B_Y F)(\chi) = F(\chi_Y). \tag{4.2}$$

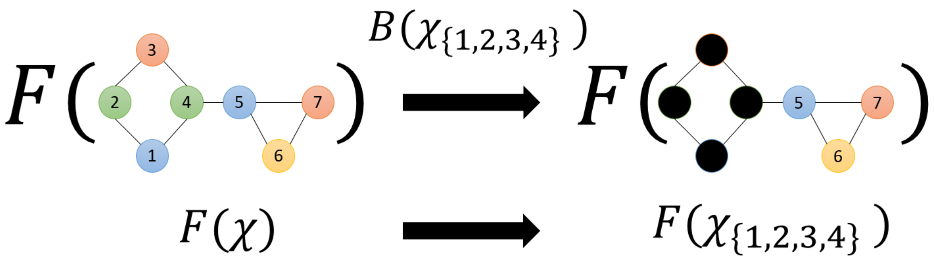

Figure 1. *Illustration of a pure blackening operator on a function F where the set of nodes $\{1,2,3,4\}$ are blackened.*

Lemma 4.4. *Let W, X and Y be three subsets of Z. The pure blackening operators B_W, B_X, B_Y have the following properties*

- *Complement: $(\mathbb{1} - B_X)$ is the unique operator B satisfying*

$$B_X \vee B = \mathbb{1} \quad \textit{and} \quad B_X \wedge B = 0. \tag{4.3}$$

We will denote the complement of an operator B as $\neg B = (\mathbb{1} - B)$. We have the two De Morgan's laws:

$$\neg(B_X \vee B_Y) = \neg B_X \wedge \neg B_Y \quad \text{and} \quad \neg(B_X \wedge B_Y) = \neg B_X \vee \neg B_Y.$$

- *Commutativity:*
  - *For $\wedge$:*

$$B_W \wedge B_X = B_X \wedge B_W \quad \text{and} \quad B_W \wedge \neg B_X = \neg B_X \wedge B_W,$$

  - *For $\vee$:*

$$B_W \vee B_X = B_X \vee B_W \quad \text{and} \quad B_W \vee \neg B_X = \neg B_X \vee B_W.$$

- *Associativity:*

$$(B_W \wedge B_X) \wedge B_Y = B_W \wedge (B_X \wedge B_Y) \quad \text{and} \quad (B_W \vee B_X) \vee B_Y = B_W \vee (B_X \vee B_Y).$$

- *Distributivity:*

$$B_W \wedge (B_X \vee B_Y) = (B_W \wedge B_X) \vee (B_W \vee B_Y).$$

The proof of this lemma can be found in Appendix B. We now introduce monomial blackening operators and polynomial blackening operators.

Definition 4.5. *Let $\{X_i\}_{i\in[1:n]}$ and $\{Y_j\}_{j\in[1:m]}$ be two finite sets of subsets of Z. We define the* monomial blackening operator *$M_{X,Y}$ associated with the sets $\{X_i\}_{i\in[1:n]}$ and $\{Y_j\}_{j\in[1:m]}$ as*

$$M_{X,Y} = \bigwedge_{1\le i\le n} B_{X_i} \bigwedge_{1\le j\le m} (\neg B_{Y_j}).$$

Let $\{M_{X^a,Y^a}\}_{a\in[1:\ell]}$ be a set of monomial blackening operators, we define the polynomial blackening operator *P as*

$$P = \bigvee_{1\le a\le \ell} M_{X^a,Y^a}.$$

We will denote the set of monomial blackening operators *as $\mathcal{M}$ and the set of* polynomials operators *as $\mathcal{P}$.*

Lemma 4.6. *We have the following properties :*

1. *$\mathcal{P}$ is stable by $\wedge, \vee$ and $\neg$.*
2. *The elements of $\mathcal{P}$ commutes two by two:*

$$\forall (P,Q) \in \mathcal{P}^2, \qquad P \wedge Q = Q \wedge P.$$

3. *Every element of $\mathcal{P}$ is a projector:*

$$\forall P \in \mathcal{P}, \qquad P \wedge P = P.$$

The proof of this lemma can be found in Appendix C.

Definition 4.7. *We define the relation $\le_{\mathcal{P}}$ on the set $\mathcal{P}$ as follows.*

$$For\,all\, P, Q \in \mathcal{P}^2, P \le_{\mathcal{P}} Q \quad \text{if, and only if,} \quad P \wedge Q = Q \quad \text{which is equivalent to} \quad P \vee Q = P$$

Lemma 4.8. *The set $(\mathcal{P}, \leq_{\mathcal{P}})$ is a partially ordered set and the operations $\vee$ and $\wedge$ in Definition 4.2 are respectively the least upper bound and the greatest lower bound as defined in Section 3. Moreover, $\neg$ is the complementary in the sense of Boolean Algebra.*

The proof of Lemma 4.8 can be found in Appendix D. In the sequel, we will denote $\leq_{\mathcal{P}}$ as $\leq$ without ambiguity.

Proposition 4.9. *The set of polynomial blackening operators $\mathcal{P}$ with the relation $\wedge$, $\vee$ and $\neg$ is a Boolean Algebra. It is the smallest Boolean Algebra containing the pure blackening operators. Moreover B_Z is the neutral element of $\mathcal{P}$ for the addition and $B_\emptyset$ is the identity element of $\mathcal{P}$ (neutral element for $\wedge$).*

The proof of Proposition 4.9 can be found in Appendix E. We will now simplify Corollary 3.11 when we consider the union and intersection of pure blackening operators.

Corollary 4.10. *Let $\{X_k\}$ and $\{Y_k\}$ be two set of subsets of Z and denote $a_k = B_{X_k}$ and $b_k = B_{Y_k}$. We can simplify the 6 conditions of Corollary 3.11 as follows:*
1. $\forall (k,j) \in (J \setminus K)^2, \quad X_j \not\subset Y_k.$
2. $\forall j \in J \setminus K,\ \forall i \subset K, \quad X_j \not\subset X_i.$
3. $\forall k \in J \setminus K,\ \forall i \subset K, \quad X_i \cup Y_k \neq Z.$
4. $\forall (k,k') \in (J \setminus K)^2, \quad Y_k \cup Y_{k'} \neq Z.$
5. $\forall (i,i') \in K^2, \quad X_i \cup X_{i'} \neq Z.$
6. $\forall (j,j') \in (J \setminus K)^2, \quad (\mathbb{1} - B_{X_j})(\mathbb{1} - B_{X_{j'}}) \neq 0.$

The proof of this corollary can be found in Appendix F. We have now the tool to study the reduction of principal ideals associated to the Boolean Algebra of the polynomial blackening operators.

5. Study of Markovian relations.

The objective of this section is to provide another proof of Hammersley-Clifford's Theorem (6) using the theory of ideals of Boolean rings.

Definition 5.1. *Let χ be a coloration of the graph G. We denote by the $\mathbb{P}(\omega = \chi)$ (in short $\mathbb{P}(\chi)$) the probability that a random coloring ω matches χ. Moreover, for all $Y \subset Z$, we denote by $\mathbb{P}(\chi^Y)$ the probability that the restriction of ω on Y matches the restriction of χ on Y. With the notation $C_{X,\chi} = \{\xi \in C \mid \forall z_i \in X, \xi(z_i) = \chi(z_i)\}$,*

$$\mathbb{P}(\chi^X) = \mathbb{P}(\omega|_X = \chi|_X) = \sum_{\xi \in C_{X,\chi}} \mathbb{P}(\xi).$$

Furthermore, for all $X \subset Z$, $Y \subset Z$, $\mathbb{P}(\chi^X, \chi^Y)$ is the probability that ω simultaneously has the partial colouring $\chi|_X$ on X as well as the partial colouring $\chi|_Y$ on Y. In addition, we note $\mathbb{P}(\chi^X|\chi^Y)$ the probability that the random colouring ω matches the specified colouring χ on the set X knowing that ω as the colouring $\chi|_Y$ on the set Y.

$$\mathbb{P}(\chi^X|\chi^Y) = \frac{\sum_{\xi \in C^{X \cup Y, \chi}} \mathbb{P}(\xi)}{\sum_{\xi \in C^{Y,\chi}} \mathbb{P}(\xi)} = \frac{\mathbb{P}(\chi^{X \cup Y})}{\mathbb{P}(\chi^Y)}. \tag{5.1}$$

Definition 5.2. *A random variable is said to be* globally Markovian *if it is Markovian for every subsets of Z. With the notation previously introduced, the Markovian assumption can be formulated as*

$$\forall X \subset Z, \quad \mathbb{P}(\chi^X|\chi^{Z\setminus X}) = \mathbb{P}(\chi^X|\chi^{\partial X}), \tag{5.2}$$

where ∂X is the set of all the neighbours of the nodes of X.

$$\partial X = \{z_j \in Z\setminus X \mid \exists z_i \in X, \quad (z_i, z_j) \in E \quad \text{or} \quad (z_j, z_i) \in E\},$$

where E is the set of edges of G.

Definition 5.3. *We define a* clique *of a graph as a set of nodes where each node is neighbour of each other. We denote by L the set of cliques of G. Moreover, given a coloration χ of the graph, we define a* light clique *as a clique where every node is not black. We denote this set $L(\chi)$.*

Theorem 5.4 (Hammersley-Clifford's Theorem (6)). *Let ω be a random coloring of Z which follows the globally Markovian properties. Suppose that $\mathbb{P}(\omega = \chi_Z) \neq 0$. Then there exist $S \in \mathcal{F}$ such that we can factorize the probability as follow:*

$$\mathbb{P}(\chi) = \mathbb{P}(\chi_Z) \exp\left(\sum_{Y \subset L(\chi)} S(\chi_{Z\setminus Y})\right). \tag{5.3}$$

Our reformulation of Theorem 5.4 consists in identifying a specific principal ideal in the Boolean ring of Blackening operators and reducing it using Corollary 4.10. Let us introduce the specific principal ideal $I(\beta)$.

Definition 5.5. *Let Z be the set of nodes of a graph G and for all node $z_i \in Z$, let ∂z_i the set of all its neighbours. We define the polynomial operator β_i as follow :*

$$\beta_i = B_{z_i} \vee B_{Z\setminus\{z_i \cup \partial z_i\}}.$$

For $X \subset Z$, we define

$$\beta_X = B_X \vee B_{Z\setminus\{X \cup \partial X\}}.$$

And we also define β as the product of all the β_i,

$$\beta = \prod_{z_i \in Z} \beta_i = \prod_{z_i \in Z} (B_{z_i} \vee B_{Z\setminus(z_i \cup \partial z_i)}).$$

Let us state two technical lemmas. We first reduce the formulation of $I(\beta)$ using Corollary 4.10.

Lemma 5.6. *$I(\beta)$ can be decomposed as :*

$$I(\beta) = \sum_{X \in L(\chi) \cup \{\emptyset\}} I(B_{Z\setminus X}), \tag{5.4}$$

where $L(\chi)$ is the set of light cliques of the graph G associated to the coloration χ.

The proof of this lemma can be found in Appendix G.

Lemma 5.7. *Let χ be a coloration of the set of nodes Z. Suppose that ω is globally Markovian. For all $X \subset Z$, we introduce the function*

$$Q_X = \log(\mathbb{P}(\chi^X)).$$

Then we have

$$\forall X \subset Z, \quad Q_Z = \beta_X Q_Z.$$

The proof of this lemma can be found in Appendix H. Using Lemma 5.6 and Lemma 5.7, we are now able to provide a new proof of Hammersley-Clifford's Theorem.

Proof of Theorem 5.4 . Let G be a graph and χ be a coloration on the graph G. Let ω be a random coloring of Z which follows the globally Markovian properties. Suppose that $\mathbb{P}(\omega = \chi_Z) \neq 0$. Using Lemma 5.7 and remarking that $Q_Z = \log(\mathbb{P}(\chi))$, we get

$$\forall z_i \in Z, \quad \beta_i \log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi)).$$

Hence,

$$\beta \log(\mathbb{P}(\chi)) = \prod_{z_i \in Z} \beta_i \log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi)).$$

Thus, $\beta \log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi))$.

Moreover, by Lemma 5.6, $\beta \in I(\beta) = \sum_{X \in L(\chi) \cup \{\emptyset\}} I(B_{Z\setminus X})$. Thus, there exists a set of projectors $E_{Z\setminus X}$ such that $E_{Z\setminus X} \in I(B_{Z\setminus X})$ and

$$\beta = \sum_{X \in L(\chi) \cup \{\emptyset\}} E_{Z\setminus X} = \sum_{X \in L(\chi) \cup \{\emptyset\}} B_{Z\setminus X} E_{Z\setminus X}, \tag{5.5}$$

$$\begin{aligned} \log(\mathbb{P}(\chi)) = \beta \log(\mathbb{P}(\chi)) &= \sum_{X \in L(\chi) \cup \{\emptyset\}} B_{Z\setminus X} E_{Z\setminus X} \log(\mathbb{P}(\chi)), \\ &= \sum_{X \in L(\chi) \cup \{\emptyset\}} E_{Z\setminus X} \log(\mathbb{P}(\chi_{Z\setminus X})). \end{aligned} \tag{5.6}$$

Finally, as $\mathbb{P}(\omega = \chi_Z) \neq 0$, with the notation $S(\chi_{Z\setminus X}) = E_{Z\setminus X} \log(\mathbb{P}(\chi_{Z\setminus X}))$, we have the final result:

$$\mathbb{P}(\chi) = \mathbb{P}(\chi_Z) \exp\Big(\sum_{X \in L(\chi)} S(\chi_{Z\setminus X})\Big). \tag{5.7}$$

∎

6. General strategy for reduction of probability laws on graph under constraints. In this section, we present a strategy to reduce deep learning models on graphs. We first sketch the main steps of this strategy, then we explain why this method works by explaining the link between Propositional Logic and random variables respecting dependency constraints and finally we apply this method to tree structure dependencies.

6.1. Description of the strategy. Let us generalize the methodology followed in Section 5 in order to find a minimization formula for a probability function under constraints on a graph. The notion of dependency between nodes is essential in this methodology. We shall say that a node z_i is independent of another node z_j if, and only if, z_j can be blackened independently of z_i. The strategy can be summarized as follows.

1. **Step 1: Identification of the invariance properties.** Identify the dependency relationships between the nodes in terms of local invariance properties. The goal of this first step is to identify locally which nodes are independent of which other nodes.
2. **Step 2: Construction of the associated blackening operators.** For each node z_i of the graph G, formulate the invariance properties identified in Step 1 in terms of invariance under a specific blackening operator β.
3. **Step 3: Reduction of the blackening operators.** Use a result similar to Corollary 4.10 to find the reduced form of the principal ideal generated by β.
4. **Step 4: Link to probability function.** In order to conclude, prove that for a random coloration ω satisfying the invariance properties identified in Step 1,

$$\beta \log(\mathbb{P}) = \log(\mathbb{P}),$$

where β has been defined in Step 2.

In the next section, we examplarize our strategy on a tree structure with non-symmetric dependency relations between nodes.

6.2. Link with Propositional logic. In this section we highlight an isomorphism between the Blackening Algebra studied above and the two-element Boolean Algebra associated with the fact that a function is invariant or not to some operators.

Definition 6.1. *We call the* two-element Boolean Algebra *the Boolean Algebra*

$$(\{\text{True}, \text{False}\}, \cup, \cap, \text{False}, \text{True}),$$

where the operations are defined in the truth table 1.

We now define the propositional function (in the sense of Propositional Logic) related to the invariance properties of element of $\mathcal{F}$. The notations $\mathcal{F}$ and $\mathcal{P}$ were introduced in Section 4.

Definition 6.2. *For all function $F \in \mathcal{F}$, we define ψ_F as the application from $\mathcal{P}$ to* $\{\text{True}, \text{False}\}$ *such that*

$$\begin{aligned} \psi_F \colon \mathcal{P} &\to \{\text{True}, \text{False}\} \\ P \to \psi_F(P) &= \begin{cases} \text{True} & \text{if} \quad \text{PF} = \text{F}, \\ \text{False} & \text{otherwise}. \end{cases} \end{aligned} \tag{6.1}$$

We will now study the function ψ_F.

Proposition 6.3. *For all $F \in \mathcal{F}$, ψ_F is a morphism in the following sense. For all $(P, Q) \in \mathcal{P}^2$, we have*

x	y	$x \cap y$	$x \cup y$
False	False	False	False
True	False	False	True
False	True	False	True
True	True	True	True

Table 1
Truth table for binary operators.

1. $$\psi_F(B_\emptyset) = \text{True},$$
2. $$\psi_F(B_Z) = \text{False},$$
3. $$\psi_F(P \vee Q) = \psi_F(P) \cup \psi_F(Q),$$
4. $$\psi_F(P \wedge Q) = \psi_F(P) \cap \psi_F(Q),$$
5. $$\psi_F(\neg P) = \neg \psi_F(P).$$

The proof of Proposition 6.3 can be found in Appendix I. This result explains why in Step 2 the construction of β_i based on two polynomial operators and the definition of β as the product of the β_i was fruitful.

6.3. Example of application: tree structure dependencies. In the field of Natural Language Processing (NLP), one of the challenges is to accurately identify the labels of words in a phrase using Artificial Intelligence algorithms. When training on small sets of examples, the results can be unsatisfying. One solution to improve the performance is to impose constraints on the consistency of the labels assigned to the words in the sentence. In this section, we propose a novel approach to tackle this problem by using the Constituency Tree representation of the sentence and imposing a consistency property on the labels of the nodes of the tree. We will explain this concept in detail in the following subsections.

After introducing some key definitions, we will demonstrate how to apply the method outlined in Subsection 6.1 to determine the minimal expression of the joint probability of the nodes of the tree when the random variables conform to the consistency property.

Definition 6.4. *An* arborescence *is a directed graph G in which, for a specific node u (called the root) and any other node v, there is exactly one directed path (i.e. a sequence of edges) from u to v . We can view an arborescence as a directed rooted tree.*

Within an arborescence, we can introduce the concepts of children, parents and siblings of a node.

Definition 6.5. *The* children *(*parents*) of a node z_i is the set of nodes $C(z_i)$ (resp. $P(z_i)$) composed of all the nodes z_j where (z_i, z_j) (resp. (z_j, z_i)) is an edge of the arborescence.*

$$C(z_i) = \{z_j | (z_i, z_j) \in E\}, \quad P(z_i) = \{z_j | (z_j, z_i) \in E\}.$$

The descendants *$D(z_i)$ (resp.* ancestors *$A(z_i)$) of z_i are defined as the sets of the children (resp. parents) of z_i and the children (resp. parents) of its children (resp. parents) recursively. The* siblings *$Sib(z_i)$ of z_i are defined as the set of nodes which have the same parents as z_i.*

We extend this notions to any subset Y of the arborescence:

$$D(Y) = \left(\bigcup_{z_i \in Y} D(z_i) \right) \setminus Y, \quad A(Y) = \left(\bigcup_{z_i \in Y} A(z_i) \right) \setminus Y, \quad Sib(Y) = \left(\bigcup_{z_i \in Y} Sib(z_i) \right) \setminus Y.$$

One important notion that we will use in the sequel is the filter on an arborescence.

Definition 6.6. *A subset F of a partially ordered set Q is an ordered filter if the following conditions hold:*

- *F is non-empty.*
- *F is downward directed: for every $x, y \in F$, there is some $z \in F$ such that $z \leq x$ and $z \leq y$.*
- *F is an upper set: for every $x \in F$, $\quad p \in Q, \quad x \leq p$ implies that $p \in F$.*

In the case of a tree structure, every set of neighbouring leaves with their common ancestors is a filter for the order induced by the direction of the edges. We will denote by $F(G)$ the set of filters of the graph G.

Figure 2 compares the factorized structure of cliques in graphs and filter in trees.

Step 1: Identification of the invariance properties..

Definition 6.7. *Let χ be a coloration on the arborescence A. Then χ is said to have the blackening consistency property if*

$$\forall z \in A, \quad \chi(z) = \text{black} \quad \textit{if and only if} \quad \forall z_j \in C(z), \quad \chi(z_j) = \text{black}.$$

As illustrated in Figure 3, this rule propagates the blackening color between branch nodes and leaf nodes. Under this constraint, any node z_i depends on the set of nodes $\{Sib(z_i) \cup P(z_i)\}$. Moreover, if a node z_i does not depends on z_j (i.e. we can blacken z_j without changing the coloration of z_i) then it also does not depend on $C(z_j)$. More generally, the set $z_i \cup Sib(z_i) \cup A(z_i) \cup D(z_i)$ and its complementary are independent.

Definition 6.8. *Let χ be a coloration on the arborescence A. A random variable is said to be* blackening consistent *if it respects the following property:*

$$(6.2) \qquad \forall X \subset Z, \quad \mathbb{P}\left(\chi^{X \cup D(X)} | \chi^{Z \setminus (X \cup D(X))}\right) = \mathbb{P}\left(\chi^{X \cup D(X)} | \chi^{Sib(X) \cup P(X)}\right).$$

Let us state the main result of this section.

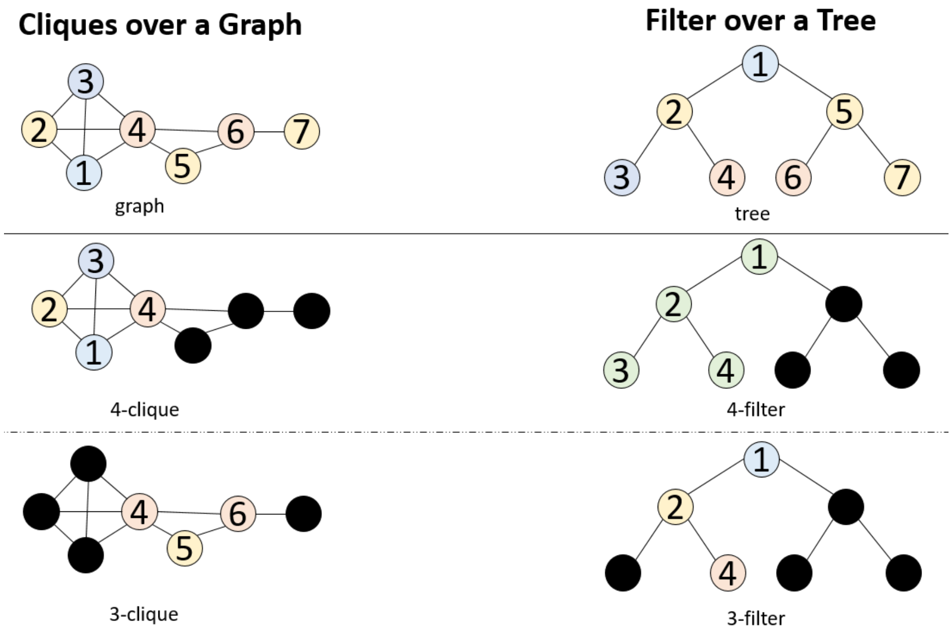

Figure 2. *Illustration of the reduction structure on graph and tree.*

Theorem 6.9. *Let ω be a random coloring of Z which follows the blackening consistent property. Suppose that $\mathbb{P}(\omega = \chi_Z) \neq 0$. We can factorize the probability as follows:*

$$\exists S \in \mathcal{F}, \quad \mathbb{P}(\chi) = \mathbb{P}(\chi_Z)\exp\left(\sum_{Y\subset F(G)} S(\chi_Y)\right). \tag{6.3}$$

In other words, the probability law of colorations can be factorized on the filters of the arborescence.

Proof. To prove this theorem, we will follow the steps 2, 3 and 4 presented in Subsection 6.1.

Step 2: Construction of the associated blackening operators. We now want to construct the blackening operator associated to the blackening consistency property. Consider a coloration χ which satisfies the blackening consistency property. Let $z_j \in Z$. We analyze the invariance of $\log(\mathbb{P}(\chi^{z_j}))$ under some blackening operators.

Let z_i be another node of the arborescence. Using the analysis of Step 1, we have only two possibilities

- Case 1: z_j is independent of z_i, the prediction does not depends on z_i and thus the log probability is invariant under the action of the operator $B_{z_i\cup D(z_i)}$.

  $$B_{z_i\cup D(z_i)}\log(\mathbb{P}(\chi^{z_j}) = \log(\mathbb{P}(\chi^{z_j}),$$

  which can be reformulated using the notation of Propositional Logic as

  $$\psi_{\log(\mathbb{P}(\chi^{z_j})}(B_{z_i\cup D(z_i)}) = \text{True}.$$

- Case 2: z_j depends on z_i, thus $z_j \in \{z_i \cup A(z_i) \cup D(z_i) \cup Sib(z_i)\}$. As the set $\{z_i \cup A(z_i) \cup D(z_i) \cup Sib(z_i)\}$ is independent of its complementary, the log probability is invariant under the action of the operator $B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))}$, i.e.

$$B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))} \log(\mathbb{P}(\chi^{z_j}) = \log(\mathbb{P}(\chi^{z_j}),$$

  which can be reformulated using the notation of Propositional Logic as

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))}) = \text{True}.$$

This two cases represent all the possibilities, thus

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(B_{z_i\cup D(z_i)}) \cup \psi_{\log(\mathbb{P}(\chi^{z_j})}(B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))}) = \text{True}.$$

Hence, using the link between the union in Propositional logic and the union in Boolean algebra presented in Proposition 6.3, we deduce that

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(B_{z_i\cup D(z_i)} \vee B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))}) = \text{True}.$$

To simplify the notations, we can introduce β_{z_i} as

$$\beta_{z_i} = B_{Z\setminus(z_i\cup A(z_i)\cup D(z_i)\cup Sib(z_i))} \vee B_{z_i\cup D(z_i)}. \tag{6.4}$$

Then

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(\beta_{z_i}) = \text{True}.$$

As this result is true for all $z_i \in Z$, we have

$$\bigcap_{z_i\in Z} \psi_{\log(\mathbb{P}(\chi^{z_j})}(\beta_{z_i}) = \text{True}.$$

Using again the link between the intersection in Propositional logic and the intersection in Boolean algebra presented in Proposition 6.3, we get

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(\prod_{z_i\in Z} \beta_{z_i}) = \text{True}.$$

To simplify the notations, we can introduce

$$\beta = \prod_{z_i\in Z} \beta_{z_i}. \tag{6.5}$$

Then

$$\psi_{\log(\mathbb{P}(\chi^{z_j})}(\beta) = \text{True}.$$

More generally, we will prove that $\log(P(\chi))$ is invariant under β in Step 4.

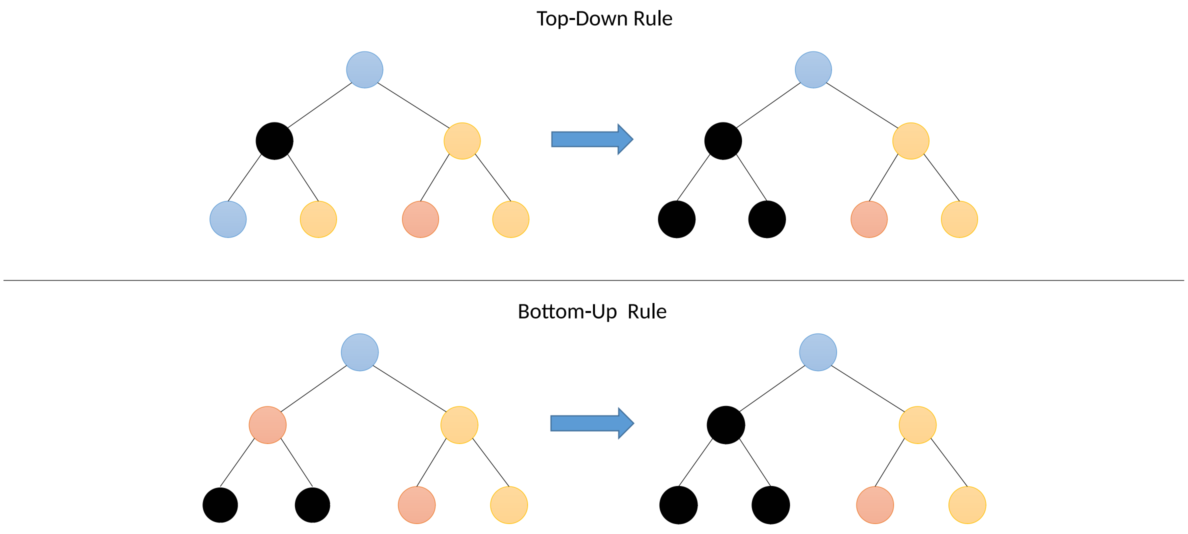

Figure 3. *Illustration of Top-Down and Bottom-Up rules. Left : inconsistent tree, right: consistent tree.*

Step 3: Reduction of the blackening operators. We can now formulate the factorization formulation of β in the following lemma.

Lemma 6.10. *$I(\beta)$ can be decomposed as*

$$I(\beta) = \sum_{I \in F(G)} I(B_I), \tag{6.6}$$

where $F(G)$ is the set of all the sets of leaves with their common ancestors.

The Proof of Lemma 6.10 can be found in Appendix J.

Step 4: Link with probability function. The last step of our methodology is proof of Theorem 6.9 can be found in Appendix K.

7. Conclusion and Future Works. In this paper, we generalized the Hammersley-Clifford theorem (6) to principal ideals on Boolean rings. This allowed to identify relations between blackening operators and Boolean algebra. We then proposed a new method to analyse data structure and nodes relationship. Finally, we illustrated this method on a specific tree structure.

In a future work, we will develop Graph Neural Networks based on the method presented in Section 6.

Appendix A. Proof of Corollary 3.11.

Proof of Corollary 3.11. With the notations introduced in Corollary 3.11, we denote by Γ the product of ideals

$$\Gamma = \left(\bigcap_{k_1 \in K} I(a_{k_1}) \bigcap_{k_2 \in J \setminus K} I(b_{k_2}) \bigcap_{k_3 \in J \setminus K} I(a'_{k_3}) \right).$$

In order to prove Corollary 3.11, we will first recall the link between orthogonality of principal ideals and orthogonality of the elements which generate them ((20) Theorem 19)) in Lemma A.1. Then, we will use it to identify a set of sufficient conditions under which the set Γ is not empty.

Lemma A.1. *Let a and b be two elements of A, the principal ideals $I(a)$ and $I(b)$ are orthogonal if and only if a and b are orthogonal.*

A sufficient condition for Γ to be null is that at least two of the ideals composing its product are orthogonal. Thus, in order that Γ to be not null, it has to verify the following 6 conditions:

1. $\forall (k_1, k'_1) \in K^2, \quad I(a_{k_1}) \not\perp I(a_{k'_1}),$
2. $\forall (k_2, k'_2) \in (J \setminus K)^2, \quad I(b_{k_2}) \not\perp I(b_{k'_2}),$
3. $\forall (k_3, k'_3) \in (J \setminus K)^2, \quad I(a'_{k_3}) \not\perp I(a'_{k'_3}),$
4. $\forall k_1 \in K, \forall k_2 \in J \setminus K, \quad I(a_{k_1}) \not\perp I(b_{k_2}),$
5. $\forall k_1 \in K, \forall k_3 \in J \setminus K, \quad I(a_{k_1}) \not\perp I(a'_{k_3}),$
6. $\forall (k_2, k_3) \in (J \setminus K)^2, \quad I(b_{k_2}) \not\perp I(a'_{k_3}).$

By using Lemma A.1, these conditions are equivalent to conditions on the generators of the principal ideals:

1. $\forall (k_1, k'_1) \in K^2, \quad a_{k_1} a_{k'_1} \neq 0,$
2. $\forall (k_2, k'_2) \in (J \setminus K)^2, \quad b_{k_2} b_{k'_2} \neq 0,$
3. $\forall (k_3, k'_3) \in (J \setminus K)^2, \quad a'_{k_3} a'_{k'_3} \neq 0,$
4. $\forall k_1 \in K, \forall k_2 \in J \setminus K, \quad a_{k_1} b_{k_2} \neq 0,$
5. $\forall k_1 \in K, \forall k_3 \in J \setminus K, \quad a_{k_1} a'_{k_3} \neq 0,$
6. $\forall (k_2, k_3) \in (J \setminus K)^2, \quad b_{k_2} a'_{k_3} \neq 0,$

Since $a'_k = \mathbb{1} \oplus a_k$, we can reformulate the conditions 5 and 6:

1. $\forall k_1 \in K, \forall k_3 \in J \setminus K, a_{k_3} a_{k_1} \neq a_{k_1}$ *i.e.* $a_{k_1} \not\leq a_{k_3}$,
2. $\forall (k_2, k_3) \in (J \setminus K)^2, \quad a_{k_3} b_{k_2} \neq b_{k_2}$ *i.e.* $b_{k_2} \not\leq a_{k_3}$,

which yields the result. ∎

Appendix B. Proof of Lemma 4.4.

Proof of Lemma 4.4. Let W, X, Y be three subsets of Z, let F be a function of $\mathcal{F}$ and χ be a coloration.

- Complementarity.
  We will first prove that $\mathbb{1} - B_X$ is a solution of Equation (4.3). We have indeed

$$\begin{aligned} B_X \wedge (\mathbb{1} - B_X) &= B_X - B_X \wedge B_X \\ &= B_X - B_X \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} B_X \vee (\mathbb{1} - B_X) &= B_X + (\mathbb{1} - B_X) - B_X \wedge (\mathbb{1} - B_X) \\ &= B_X + (\mathbb{1} - B_X) - 0 \\ &= \mathbb{1}. \end{aligned}$$

Let us show that it is the unique solution of Equation (4.3).
Let B be a solution of Equation (4.3). Then, we have

$$B_X \wedge B = 0 \quad \text{and} \quad B_X \vee B = B_X + B - B_X \wedge B = \mathbb{1}.$$

By inserting the first equation into the second, we obtain the system

$$B_X \wedge B = 0 \quad \text{and} \quad B_X + B = \mathbb{1}.$$

Therefore $B = \mathbb{1} - B_X$.
Let us prove the two De Morgan's laws. We have

$$\begin{aligned} \neg(B_X \vee B_Y) &= \mathbb{1} - B_X \vee B_Y \\ &= \mathbb{1} - (B_X + B_Y - B_X \wedge B_Y) \\ &= (\mathbb{1} - B_X) - B_Y \wedge (\mathbb{1} - B_X) \\ &= (\mathbb{1} - B_X) \wedge (\mathbb{1} - B_Y) \\ &= \neg B_X \wedge \neg B_Y \end{aligned}$$

and

$$\begin{aligned} \neg(B_X \wedge B_Y) &= \mathbb{1} - B_X \wedge B_Y \\ &= \mathbb{1} + (\mathbb{1} - \mathbb{1}) + (B_X - B_X) + (B_Y - B_Y) - B_X \wedge B_Y \\ &= (\mathbb{1} - B_X) + (\mathbb{1} - B_Y) - (\mathbb{1} - B_X - B_Y + B_X \wedge B_Y) \\ &= (\mathbb{1} - B_X) + (\mathbb{1} - B_Y) - (\mathbb{1} - B_X) \wedge (\mathbb{1} - B_Y) \\ &= \neg B_X \vee \neg B_Y. \end{aligned}$$

- Commutativity
  - We first prove the commutativity of the operator $\wedge$:

$$\begin{aligned} (B_W \wedge B_X) F(\chi) &= B_W F(\chi_X) = F(\chi_{X \cup W}) \\ &= B_X F(\chi_W) = (B_X \wedge B_W) F(\chi) \end{aligned}$$

  and

$$\begin{aligned} B_W \wedge \neg B_X &= B_W \wedge (\mathbb{1} - B_X) = B_W - B_W \wedge B_X \\ &= B_W - B_X \wedge B_W = (\neg B_X \wedge B_W). \end{aligned}$$

  - We now prove the commutativity of the operator $\vee$:

$$\begin{aligned} B_W \vee B_X &= B_W + B_X - B_W \wedge B_X = B_X + B_W - B_W \wedge B_X \\ &= B_X + B_W - B_X \wedge B_W = B_X \vee B_W \end{aligned}$$

and

$$\begin{aligned} B_W \vee \neg B_X &= B_W + \neg B_X - B_W \wedge \neg B_X = \neg B_X + B_W - B_W \wedge \neg B_X \\ &= \neg B_X + B_W - \neg B_X \wedge B_W = \neg B_X \vee B_W. \end{aligned}$$

- Associativity
  - We first prove the associativity of the operator $\wedge$:

$$\begin{aligned} ((B_W \wedge B_X) \wedge B_Y)F(\chi) &= (B_W \wedge B_X)F(\chi_Y) = F(\chi_{W\cup X\cup Y}) \\ &= B_W \wedge F(\chi_{X\cup Y}) = B_W \wedge (B_X \wedge B_Y) \circ F(\chi). \end{aligned}$$

  - We now prove the associativity of the operator $\vee$:

$$\begin{aligned} (B_W \vee B_X) \vee B_Y &= (B_W \vee B_X) + B_Y - (B_W \vee B_X) \wedge B_Y \\ &= (B_W + B_X - B_{W\cup X}) + B_Y - (B_W + B_X - B_{W\cup X}) \wedge B_Y \\ &= B_W + B_X + B_Y - B_{W\cup X} - B_{W\cup Y} - B_{X\cup Y} + B_{W\cup X\cup Y} \\ &= (B_X + B_Y - B_{X\cup Y}) + B_W - B_W \wedge (B_X + B_Y - B_{X\cup Y}) \\ &= B_W \vee (B_X \vee B_Y). \end{aligned}$$

- Distributivity
  - We prove the distributivity the $\wedge$ over $\vee$:

$$\begin{aligned} B_W \wedge (B_X \vee B_Y) &= B_W \wedge (B_X + B_Y - B_{X\cup Y}) \\ &= B_W \wedge B_X + B_W \wedge B_Y - B_{W\cup X\cup Y} \\ &= B_W \wedge B_X + B_W \wedge B_Y - B_{W\cup X} \wedge B_{W\cup Y} \\ &= (B_W \wedge B_X) \vee (B_W \wedge B_Y). \end{aligned}$$ ∎

Appendix C. Proof of Lemma 4.6.

Proof of Lemma 4.6. 1. Let P be a polynomial operator with the same notation as in Definition 4.5. By Lemma 4.4,

$$\begin{aligned} \neg P &= \neg\big(\bigvee_{1\le a\le l} M_{X_a,Y_a}\big) \\ &= \bigwedge_{1\le a\le l} (\neg M_{X_a,Y_a}) \\ &= \bigwedge_{1\le a\le l} \big(\neg\big(\bigwedge_{1\le i\le n} B_{X_{a,i}} \bigwedge_{1\le j\le m} (\neg B_{Y_{a,j}})\big)\big) \\ &= \bigwedge_{1\le a\le l} \big(\bigvee_{1\le i\le n} (\neg B_{X_{a,i}}) \bigvee_{1\le j\le m} B_{Y_{a,j}}\big). \end{aligned}$$

This formulation can be rewritten as a sum of products because of the distributivity of $\wedge$ over $\vee$. Thus, $\neg P$ is a polynomial operator.
The stability by $\vee$ is obvious and the stability by $\wedge$ comes from the distributivity of $\wedge$ over $\vee$.

2. The fact that the polynomial operators commute comes again from the distributivity of $\wedge$ over $\vee$ and the fact that pure blackening operators commute.
3. Let P and Q be two elements of $\mathcal{P}$ which are projectors. Since P and Q commute, we have

$$(P \wedge Q)^2 = (P \wedge Q) \wedge (P \wedge Q) = (P \wedge P) \wedge (Q \wedge Q) = P \wedge Q,$$

$$\begin{aligned}(P \vee Q)^2 &= (P \vee Q) \wedge (P \vee Q)\\ &= (P + Q - P \wedge Q) \wedge (P + Q - P \wedge Q)\\ &= (P^2 + P \wedge Q - P \wedge Q) + (P \wedge Q + Q^2 - P \wedge Q) - (P \wedge Q)^2\\ &= P + Q - P \wedge Q = P \vee Q,\end{aligned}$$

$$(\neg P)^2 = (\mathbb{1} - P) \wedge (\mathbb{1} - P), = (\mathbb{1} - P) - P + P^2 = \mathbb{1} - P = \neg P.$$

Now we notice that pure blackening operators are projectors. Hence, by direct induction, monomial blackening operators are projectors, then polynomial blackening operator are also projectors. ∎

Appendix D. Proof of Lemma 4.8.

Proof of Lemma 4.8. Let P be an element of $\mathcal{P}$, by Item 3 of Lemma 4.6,

$$P \wedge P = P.$$

Thus, $P \leq_{\mathcal{P}} P$ and $\leq_{\mathcal{P}}$ is reflexive.

Let P and Q be two elements of $\mathcal{P}$ such that $P \leq_{\mathcal{P}} Q$ and $Q \leq_{\mathcal{P}} P$. Then, we have

$$P \wedge Q = P \quad \textit{and} \quad Q \wedge P = Q.$$

Item 2 of Lemma 4.6 implies that $P = Q$. Therefore $\leq_{\mathcal{P}}$ is anti-symmetric.

Let P, Q and R be three elements of $\mathcal{P}$ such that $P \leq_{\mathcal{P}} Q$ and $Q \leq_{\mathcal{P}} R$. Then,

$$P \wedge R = P \wedge (Q \wedge R) = (P \wedge Q) \wedge R = Q \wedge R = R.$$

This proves that $P \leq_{\mathcal{P}} R$ and that $\leq_{\mathcal{P}}$ is transitive.

We have proved that $\leq_{\mathcal{P}}$ is a partial order. We will now prove that $P \vee Q$ is the greatest lower bound of the set $\{P, Q\}$. Let P, Q be two elements of $\mathcal{P}$, we have

$$(P \vee Q) \wedge P = P \wedge P + P \wedge Q - P \wedge Q \wedge P = P$$

and

$$(P \vee Q) \wedge Q = P \wedge Q + Q \wedge Q - P \wedge Q \wedge Q = Q.$$

Thus $P \vee Q \leq_{\mathcal{P}} P$ and $P \vee Q \leq_{\mathcal{P}} Q$. This means that $P \vee Q$ is a lower bound of the set $\{P, Q\}$.

Let R such that $R \leq_{\mathcal{P}} P$ and $R \leq_{\mathcal{P}} Q$. Then, we have

$$R \wedge (P \vee Q) = (R \wedge P) + (R \wedge Q) - (R \wedge P \wedge Q) = P + Q - P \wedge Q = P \vee Q.$$

Thus, $R \leq_{\mathcal{P}} (P \vee Q)$ and $P \vee Q$ is the greatest lower bound of the set $\{P, Q\}$. We will now prove that $P \wedge Q$ is the least upper bound of the set $\{P, Q\}$.

Let P, Q and R be three elements of $\mathcal{P}$ such that $P \leq_{\mathcal{P}} R$ and $Q \leq_{\mathcal{P}} R$. We have

$$(P \wedge Q) \wedge P = P \wedge Q \quad \textit{and} \quad (P \wedge Q) \wedge Q = P \wedge Q.$$

Thus, $P \wedge Q$ is an upper bound of the set $\{P, Q\}$.

Then, we have

$$(P \wedge Q) \wedge R = P \wedge (Q \wedge R) = P \wedge R = R.$$

Thus $(P \wedge Q) \leq_{\mathcal{P}} R$ and $P \wedge Q$ is the least upper bound of the set $\{P, Q\}$.

Finally, the fact that $\neg$ is the complementary operator is direct by definition of $\neg$ and the fact that $\vee$ and $\wedge$ are the least upper bound and the greatest lower bound. ■

Appendix E. Proof of Proposition 4.9.

Proof of Proposition 4.9. As a straightforward consequence of Lemma 4.4, the lattice $(\mathcal{P}, \leq)$ is distributive and every element has a unique complement.

Let P be an operator on $\mathcal{F}$. Then,

$$\begin{aligned}(PB_Z)F(\chi) &= PF(\chi_Z) = 0\\ (PB_\emptyset)F(\chi) &= PF(\chi).\end{aligned}$$

Thus, B_Z (resp. $B_\emptyset$)is the neutral element for $\vee$ (resp. $\wedge$) on the set of the operators on $\mathcal{F}$. They are in particular the neutral elements of $(\mathcal{P}, \leq)$.

We have proved that $(\mathcal{P}, \leq)$ is complemented. It is a Boolean algebra, which concludes the proof of Proposition 4.9. ■

Appendix F. Proof of Corollary 4.10.

Proof of Corollary 4.10. Recall that $a_i = B_{X_i}$ and $b_i = B_{Y_i}$. The system of necessary conditions in order that Γ is not the empty set can be reformulated as:

1. $\forall (k_1, k_1') \in K^2, \quad B_{X_{k_1}} B_{X_{k_1'}} \neq 0 \quad \text{i.e.} \quad X_{k_1} \cup X_{k_1'} \neq Z,$
2. $\forall (k_2, k_2') \in (J \setminus K)^2, \quad B_{Y_{k_2}} B_{Y_{k_2'}} \neq 0 \quad \text{i.e.} \quad Y_{k_2} \cup Y_{k_2'} \neq Z,$
3. $\forall (k_3, k_3') \in (J \setminus K)^2, \quad (\mathbb{1} - B_{X_{k_3}})(\mathbb{1} - B_{X_{k_3'}}) \neq 0,$
4. $\forall k_1 \in K, \forall k_2 \in J \setminus K, \quad B_{X_{k_1}} B_{Y_{k_2}} \neq 0, \quad \text{i.e.} \quad X_{k_1} \cup Y_{k_2} \neq Z,$
5. $\forall k_1 \in K, \forall k_3 \in J \setminus K, \quad B_{X_{k_1}} \not\leq B_{X_{k_3}}, \quad \text{i.e.} \quad X_{k_3} \not\subset X_{k_1},$
6. $\forall (k_2, k_3) \in (J \setminus K)^2, \quad B_{Y_{k_2}} \not\leq B_{X_{k_3}} \quad \text{i.e.} \quad X_{k_3} \not\subset Y_{k_2},$

which yields the result. ■

Appendix G. Proof of Lemma 5.6.

Proof of Lemma 5.6. We can notice that $I(\beta)$ has the same form as the ideal studied in Corollary 4.10 with $\forall i \in J, X_i = z_i \, and \, Y_i = Z \backslash (z_i \cup \partial z_i)$. The system of necessary conditions in order that Γ is not null can be reformulated as:

1. $\forall (k_1, k_1') \in K, \quad z_{k_1} \cup z_{k_1'} \neq Z,$
2. $\forall (k_2, k_2') \in J \setminus K$:
   $\{Z \backslash (z_{k_2} \cup \partial z_{k_2})\} \cup \{Z \backslash (z_{k_2'} \cup \partial z_{k_2'})\} \neq Z \quad \text{i.e.} \quad \{z_{k_2} \cup \partial z_{k_2}\} \cap \{z_{k_2'} \cup \partial z_{k_2'}\} \neq \emptyset,$
3. $\forall (k_3, k_3') \in J \setminus K, \quad (\mathbb{1} - B_{z_{k_3}})(\mathbb{1} - B_{z_{k_3'}}) \neq 0,$
4. $\forall k_1 \in K, \forall k_2 \in J \setminus K, \quad z_{k_1} \cup Z \backslash (z_{k_2} \cup \partial z_{k_2}) \neq Z, \quad \text{i.e.} \quad (z_{k_2} \cup \partial z_{k_2}) \neq z_{k_1},$
5. $\forall k_1 \in K, \forall k_3 \in J \setminus K, \quad \{z_{k_3}\} \not\subset \{z_{k_1}\} \quad \text{i.e.} \quad z_{k_3} \neq z_{k_1},$
6. $\forall (k_2, k_3) \in J \setminus K, \quad z_{k_3} \not\subset Z \backslash (z_{k_2} \cup \partial z_{k_2}) \quad \text{i.e.} \quad z_{k_3} \subset z_{k_2} \cup \partial z_{k_2}.$

If the cardinality of Z is greater than 2, the condition 1 is directly satisfied.

The condition 6 implies that every element in $J \backslash K$ are all neighbours two by two. Which means that $J \backslash K$ is a clique. Thus, $J \backslash K$ is a clique is a necessary condition in order that the element is not null.

Suppose that, $J \backslash K$ is a clique. Then, the condition 2 is verified.

The condition 5 is always verified.

Thus, we can reduce the formula of $I(\beta)$ to the sum over the cliques of the graph:

$$I(\beta) = \sum_{X \in L(Z)} I(B_{Z \backslash X}) = I(\sum_{X \in L(Z)} B_{Z \backslash X}).$$

■

Appendix H. Proof of Lemma 5.7.

Proof of Lemma 5.7. First, we can notice that when we are predicting χ^X, blackening other nodes than X have no effect. Thus, we have the equality

$$\forall X \subset Z, \forall Y \quad \text{such that} \quad X \wedge Y = \emptyset, \quad \mathbb{P}(\chi_Y^X) = \mathbb{P}(\chi^X). \tag{H.1}$$

Let $X \subset Z$. By using the Markovian assumption, we get

$$\begin{aligned} Q_Z &= \log(\mathbb{P}(\chi^Z)) = \log(\mathbb{P}(\chi^X, \chi^{Z \backslash X})) \\ &= \log(\mathbb{P}(\chi^{Z \backslash X})\mathbb{P}(\chi^X | \chi^{Z \backslash X})) = \log(\mathbb{P}(\chi^{Z \backslash X})\mathbb{P}(\chi^X | \chi^{\partial X})) \\ &= \log(\mathbb{P}(\chi^{Z \backslash X})\mathbb{P}(\chi^X, \chi^{\partial X})/\mathbb{P}(\chi^{\partial X})) = \log(\mathbb{P}(\chi^{Z \backslash X})) + \log(\mathbb{P}(\chi^{X \cup \partial X})) - \log(\mathbb{P}(\chi^{\partial X})) \\ &= Q_{Z \backslash X} + Q_{X \cup \partial X} - Q_{\partial X}. \end{aligned} \tag{H.2}$$

Let Y be another subset of Z. Applying Equation (H.1) to Equation (H.2) yields

$$B_Y Q_Z = B_Y Q_{Z \backslash X} + B_Y Q_{X \cup \partial X} - B_Y Q_{\partial X},$$

which is equivalent to

$$\frac{\mathbb{P}(\chi_Y)}{\mathbb{P}(\chi_Y^{X \cup \partial X)})} = \frac{\mathbb{P}(\chi_Y^{Z \backslash X})}{\mathbb{P}(\chi_Y^{\partial X})}.$$

Let us now prove that

$$Q_{X\cup\partial X} - B_X Q_{X\cup\partial X} = B_{Z\backslash\partial X} Q_Z - B_{Z\backslash(X\cup\partial X)} Q_Z. \tag{H.3}$$

To this aim, we denote

$$S = \frac{\mathbb{P}(\chi^{X\cup\partial X})}{\mathbb{P}(\chi_X^{X\cup\partial X})} - \frac{\mathbb{P}(\chi_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi_{Z\backslash\partial X})}.$$

From Equation (H.1) we get

$$\mathbb{P}(\chi^{X\cup\partial X}) = \mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})$$

and

$$\mathbb{P}(\chi_X^{X\cup\partial X}) = \mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash\partial X}),$$

which leads to

$$S = \frac{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash\partial X})} - \frac{\mathbb{P}(\chi_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi_{Z\backslash\partial X})}.$$

Then,

$$S\frac{\mathbb{P}(\chi_{Z\backslash\partial X})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})} = \frac{\mathbb{P}(\chi_{Z\backslash\partial X})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash\partial X})} - \frac{\mathbb{P}(\chi_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})}.$$

Using now Equation (H.2), with $Y = Z\backslash\partial X$

$$\frac{\mathbb{P}(\chi_{Z\backslash\partial X})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash\partial X})} = \frac{\mathbb{P}(\chi^{Z\backslash X}_{Z\backslash\partial X})}{\mathbb{P}(\chi^{\partial X}_{Z\backslash\partial X})} = \frac{\mathbb{P}(\chi^{Z\backslash X}_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi^{\partial X}_{Z\backslash(X\cup\partial X)})}.$$

Using again Equation (H.2), with $Y = Z\backslash(X\cup\partial X)$, gives

$$\frac{\mathbb{P}(\chi_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})} = \frac{\mathbb{P}(\chi^{Z\backslash X}_{Z\backslash(X\cup\partial X)})}{\mathbb{P}(\chi^{\partial X}_{Z\backslash(X\cup\partial X)})}.$$

Consequently, we have

$$S\frac{\mathbb{P}(\chi_{Z\backslash\partial X})}{\mathbb{P}(\chi^{X\cup\partial X}_{Z\backslash(X\cup\partial X)})} = 0,$$

and S = 0 which proves Equation (H.3).

Applying now B_X to Equation (H.2),

$$\begin{aligned} B_X Q_Z &= B_X Q_{Z\backslash X} + B_X Q_{X\cup\partial X} - B_X Q_{\partial X} \\ &= B_X \log(\mathbb{P}(\chi^{Z\backslash X})) + B_X Q_{X\cup\partial X} - B_X \log(\mathbb{P}(\chi^{\partial X})) \\ &= \log(\mathbb{P}(\chi_X^{Z\backslash X})) + B_X Q_{X\cup\partial X} - \log(\mathbb{P}(\chi_X^{\partial X})) \\ &= \log(\mathbb{P}(\chi^{Z\backslash X})) + B_X Q_{X\cup\partial X} - \log(\mathbb{P}(\chi^{\partial X})) \\ &= Q_{Z\backslash X} + B_X Q_{X\cup\partial X} - Q_{\partial X}. \end{aligned} \tag{H.4}$$

Now substracting Equation (H.2) to Equation (H.4), and using Equation (H.1)

$$\begin{aligned} Q_Z(\mathbb{1} - B_X) &= B_X Q_{X\cup\partial X} - Q_{X\cup\partial X} \\ &= B_{Z\setminus(X\cup\partial X)} Q_Z - B_{Z\setminus\partial X} Q_Z. \end{aligned}$$

This leads to

$$Q_Z = B_X Q_Z + B_{Z\setminus(X\cup\partial X)} Q_Z - B_{Z\setminus\partial X} Q_Z = \beta_X Q_Z, \qquad \blacksquare$$

which end the proof.

Appendix I. Proof of Proposition 6.3.

We first prove a useful lemma.

Lemma I.1. *Let $P \in \mathcal{P}$, then for all $F \in \mathcal{F}$, we have*

$$\psi_F(P) = \text{True} \quad \textit{if, and only if,} \quad \mathbb{1} \in I(P). \tag{I.1}$$

Proof of Lemma I.1. Let $P \in \mathcal{P}$, the notation $\mathbb{1} \in I(P)$ means that

$$P \circ \mathbb{1} = \mathbb{1}.$$

This equality has to be understood in the sense of operator on $\mathcal{F}$, this is thus equivalent to

$$\forall F \in \mathcal{F}, \quad (P \circ \mathbb{1})F = \mathbb{1}F.$$

Thus

$$PF = F \quad \text{which is equivalent to} \quad \psi_F(P) = \text{True}.$$

This ends the proof. $\blacksquare$

Proof of Proposition 6.3. We first prove Item 1 and Item 2. For all $F \in \mathcal{F}$, we have $B_\emptyset \circ F = F$ and $B_Z \circ F = 0$, so clearly $\psi_F(B_\emptyset) = \text{True}$ and $\psi_F(B_Z) = \text{False}$.

Proof of Item 3 Suppose that $\psi_F(P) = \text{True}$ then,

$$\begin{aligned} (P \vee Q)F &= PF + QF - (P \circ Q)F \\ &= F + QF - QF \\ &= F. \end{aligned}$$

As it is the same if $\psi_F(Q) = \text{True}$, we have

$$\psi_F(P) \cup \psi_F(Q) = \text{True} \quad \text{implies that} \quad \psi_F(P \vee Q) = \text{True}.$$

Suppose that

$$\psi_F(P \vee Q) = \text{True}.$$

By using Lemma I.1, we get

$$\mathbb{1} \in I(P \vee Q) = I(P) \cup I(Q).$$

Then it means, in terms of ideals,

$$\mathbb{1} \in I(P) \quad \text{or} \quad \mathbb{1} \in I(Q).$$

Therefore, again by Lemma I.1,

$$\psi_F(P) = \text{True} \quad \text{or} \quad \psi_F(Q) = \text{True},$$

which leads to

$$\psi_F(P \vee Q) = \text{True} \quad \text{if and only if} \quad \psi_F(P) \cup \psi_F(Q) = \text{True},$$

which ends the proof of Item 3.

Proof of Item 4. Suppose that

$$\psi_F(P) \cap \psi_F(Q) = \text{True}.$$

Then

$$\psi_F(P) = \text{True} \quad \text{and} \quad \psi_F(Q) = \text{True}$$

and

$$(P \wedge Q)F = P(QF) = PF = F.$$

We have proved that

$$\psi_F(P) \cap \psi_F(Q) = \text{True} \quad \text{implies that} \quad \psi_F(P \wedge Q) = \text{True},$$

Suppose now that

$$\psi_F(P \wedge Q) = \text{True},$$

then by using Lemma I.1,

$$\mathbb{1} \in I(P \wedge Q) = I(P) \cap I(Q).$$

Therefore,

$$\mathbb{1} \in I(P) \quad \text{and} \quad \mathbb{1} \in I(Q)$$

which means that

$$\psi_F(P) = \text{True} \quad \text{and} \quad \psi_F(Q) = \text{True}.$$

We have proved that

$$\psi_F(P \wedge Q) = \text{True} \quad \text{if and only if} \quad \psi_F(P) \cap \psi_F(Q) = \text{True},$$

which ends the proof of Item 4.

Item 5. Suppose that

$$\psi_F(\neg P) = \text{True},$$

then by using Lemma I.1 on $\neg P$,

$$(\mathbb{1} - P)F = F, \quad PF = 0.$$

Thus,

$$\psi_F(P) = \text{False}.$$

Suppose that

$$\psi_F(P) = \text{True}$$

then by using Lemma I.1,

$$PF = F, \quad (\mathbb{1} - P)F = 0.$$

Finally,

$$\psi_F(\neg P) = \text{False}.$$

We have thus proved that

$$\psi_F(\neg P) = \text{False} \quad \text{if and only if} \quad \psi_F(P) = \text{True},$$

which is equivalent to

$$\psi_F(\neg P) = \neg\psi_F(P).$$

■

Appendix J. Proof of Lemma 6.10.

Proof of Lemma 6.10. According to Corollary 4.10, an element of the sum of Lemma 6.10 is zero if one of the six following conditions is satisfied.

1. $\forall (k,j) \in (Z \setminus I)^2$,

$$\begin{aligned} z_j \cup D(z_j) &\not\subset Z \backslash \{z_k \cup Sib(z_k) \cup A(z_k) \cup D(z_k)\} \\ \text{i.e.} \quad \{z_j \cup D(z_j)\} \cap \{z_k &\cup Sib(z_k) \cup A(z_k) \cup D(z_k)\} \neq \emptyset. \end{aligned}$$

2. $\forall j \in Z \setminus I, \forall i \in I$,

$$z_j \cup D(z_j) \not\subset z_i \cup D(z_i).$$

3. $\forall k \in Z \setminus I$, $\forall j \in I$,

$$\begin{aligned} \{z_j \cup D(z_j)\} \cup Z \backslash \{z_k \cup Sib(z_k) \cup A(z_k) \cup D(z_k)\} &\neq Z \\ \text{i.e.} \quad \{z_k \cup Sib(z_k) \cup A(z_k) \cup D(z_k)\} &\not\subset \{z_j \cup D(z_j)\}. \end{aligned}$$

4. $\forall (k,k') \in (Z \setminus I)^2$,

$$\begin{aligned} Z\backslash\{z_k \cup Sib(z_k) \cup A(z_k) \cup D(z_j)\} \cup Z\backslash\{z'_k \cup Sib(z'_k) \cup A(z'_k) \cup D(z'_k)\} &\neq Z \\ \text{i.e.} \quad (z_k \cup Sib(z_k) \cup A(z_k) \cup D(z_k)) \cap (z'_k \cup Sib(z'_k) \cup A(z'_k) \cup D(z'_k)) &\neq \emptyset. \end{aligned}$$

5. $\forall (i,i') \in I^2$,

$$z_i \cup D(z_i) \cup z'_i \cup D(z'_i) \neq Z.$$

6. $\forall (j, j') \in (Z \setminus I)^2$,

$$(\mathbb{1} - B_{z_i \cup D(z_i)})(\mathbb{1} - B_{z'_j \cup D(z'_j)}) \neq 0.$$

Condition 1 implies that every node in $Z \backslash I$ has a relation among each other node's neighbour or descendant or ancestor.

Condition 2 implies that the elements of $Z \backslash I$ are not the descendant of the elements of I.

Condition 3 implies the element of I are not the ancestor of the element of $Z \backslash I$.

Condition 4 implies that every element of $Z \backslash I$ is parent, descendant or neighbour of one another of every other element of $Z \backslash I$.

Condition 5 is satisfied when $Z \backslash I \neq \emptyset$ and at least one element of $Z \backslash I$ is not a descendant of any element of I.

These conditions imply the following fact. Assume that the set $Z \backslash I$ is not totally black. Then, necessarily, $Z \backslash I$ is composed of a set of sibling leaves Sib in addition to their ascendants $D(Sib)$. In other words, it is a filter. ∎

Appendix K. Proof of Theorem 6.9.

In order to prove Theorem 6.9, we will follow the same steps as the proof of Theorem 5.4 in the Markovian setting. Let T be an arboresence and χ be a coloration on the arboresence T. Let ω be a random coloring of Z which follows the filter invariant properties. Suppose that $\mathbb{P}(\omega = \chi_Z) \neq 0$. Let us state a useful lemma, whose proof is deferred to the end of this section.

Lemma K.1. *Let χ be a coloration of the set of nodes Z and suppose that ω follows the filter invariant properties. We introduce the function*

$$\forall X \subset Z, \quad Q_X = \log(\mathbb{P}(\chi^X)).$$

Then, we have

$$\forall X \subset Z, \quad Q_Z = \beta_X Q_Z,$$

where β is defined in 6.4.

Proof of Theorem 6.9. Let $X \subset Z$, using the filter invariant properties of ω, we get

$$\begin{aligned}
Q_Z &= \log(P(\chi^Z)) = \log(P(\chi^{X \cup D(X)}, \chi^{Z \backslash \{X \cup D(X)\}})) \\
&= \log(P(\chi^{Z \backslash \{X \cup D(X)\}}) P(\chi^{X \cup D(X)} | \chi^{Z \backslash X \cup D(X)})) \\
&= \log(P(\chi^{Z \backslash \{X \cup D(X)\}}) P(\chi^{X \cup D(X)} | \chi^{Sib(X) \cup A(X)})) \\
&= \log(P(\chi^{Z \backslash \{X \cup D(X)\}}) P(\chi^{X \cup D(X)}, \chi^{Sib(X) \cup A(X)}) / P(\chi^{Sib(X) \cup A(X)})) \\
&= \log(P(\chi^{Z \backslash \{X \cup D(X)\}})) + \log(P(\chi^{X \cup Sib(X) \cup D(X) \cup A(X)})) - \log(P(\chi^{Sib(X) \cup P(X)})) \\
&= Q_{Z \backslash \{X \cup D(X)\}} + Q_{X \cup D(X) \cup Sib(X) \cup A(X)} - Q_{Sib(X) \cup A(X)}.
\end{aligned} \tag{K.1}$$

We can now use Lemma K.1 to get

$$\forall z_i \in Z, \quad \beta_{z_i} \log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi)).$$

Therefore,

$$\begin{aligned}\beta\log(\mathbb{P}(\chi)) &= \left(\prod_{z_i\in Z}\beta_{z_i}\right)\log(\mathbb{P}(\chi)) = \left(\prod_{z_i\in Z\setminus\{z_1\}}\beta_{z_i}\right)(\beta_{z_1}\log(\mathbb{P}(\chi)))\\ &= \left(\prod_{z_i\in Z\setminus\{z_1\}}\beta_{z_i}\right)\log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi)).\end{aligned}$$

Thus, $\beta\log(\mathbb{P}(\chi)) = \log(\mathbb{P}(\chi))$.

By Lemma 6.10, there exists a set of projector $E_{Z\setminus X}$ such that $E_{Z\setminus X} = B_{Z\setminus X}E_{Z\setminus X}$ and

$$\beta = \sum_{X\in F(\chi)\cup\{\emptyset\}} E_{Z\setminus X} = \sum_{X\in F(\chi)\cup\{\emptyset\}} B_{Z\setminus X}E_{Z\setminus X},$$

so

$$\begin{aligned}\log(\mathbb{P}(\chi)) = \beta\log(\mathbb{P}(\chi)) &= \sum_{X\in F(\chi)\cup\{\emptyset\}} B_{Z\setminus X}E_{Z\setminus X}\log(\mathbb{P}(\chi))\\ &= \sum_{X\in F(\chi)\cup\{\emptyset\}} E_{Z\setminus X}\log(\mathbb{P}(\chi_{Z\setminus X})).\end{aligned}$$

Finally, as $\mathbb{P}(\omega = \chi_Z) \neq 0$, with the notation $S(\chi_{Z\setminus X}) = E_{Z\setminus X}\log(\mathbb{P}(\chi_{Z\setminus X}))$, we have the final result:

$$\mathbb{P}(\chi) = \mathbb{P}(\chi_Z)\exp\left(\sum_{X\in F(\chi)} S(\chi_{Z\setminus X})\right).$$

This ends the proof of Theorem 6.9. ■

Proof of Lemma K.1. Let Y be another subset of Z. Applying Equation (H.1) to Equation (K.1) yields

$$B_YQ_Z = B_YQ_{Z\setminus\{X\cup D(X)\}} + B_YQ_{X\cup D(X)\cup Sib(X)\cup A(X)} - B_YQ_{Sib(X)\cup A(X)},$$

which is equivalent to

$$\frac{\mathbb{P}(\chi_Y)}{\mathbb{P}(\chi_Y^{X\cup Sib(X)\cup D(X)\cup A(X)})} = \frac{\mathbb{P}(\chi_Y^{Z\setminus(X\cup D(X))})}{\mathbb{P}(\chi_Y^{Sib(X)\cup A(X)})}.$$

By applying the same proof as in Lemma 5.7 for the proof of Equation (H.3), where we replace ∂X by $Sib(X)\cup A(X)$ and X by $X\cup D(X)$, we have

$$\begin{aligned}&B_{X\cup D(X)}Q_{X\cup Sib(X)\cup D(X)\cup A(X)} - Q_{X\cup Sib(X)\cup D(X)\cup A(X)}\\ &= B_{Z\setminus(X\cup Sib(X)\cup D(X)\cup A(X))}Q_Z - B_{Z\setminus(Sib(X)\cup A(X))}Q_Z.\end{aligned}$$

This equation yields

$$Q_{X\cup Sib(X)} - B_X Q_{X\cup Sib(X)} = B_{Z\setminus Sib(X)} Q_Z - B_{Z\setminus(X\cup Sib(X))} Q_Z. \tag{K.2}$$

By applying now B_X to Equation (K.1), we get

$$\begin{aligned}
B_{X\cup D(X)} Q_Z &= B_{X\cup D(X)} Q_{Z\setminus(X\cup D(X))} + B_{X\cup D(X)} Q_{X\cup Sib(X)} - B_{X\cup D(X)} Q_{Sib(X)} \\
&= B_{X\cup D(X)} \log(\mathbb{P}(\chi^{Z\setminus(X\cup D(X))})) + B_{X\cup D(X)} Q_{X\cup Sib(X)} - B_{X\cup D(X)} \log(\mathbb{P}(\chi^{Sib(X)})) \\
&= \log(\mathbb{P}(\chi^{Z\setminus(X\cup D(X))}_{X\cup D(X)})) + B_{X\cup D(X)} Q_{X\cup Sib(X)} - \log(\mathbb{P}(\chi^{Sib(X)}_X)) \\
&= \log(\mathbb{P}(\chi^{Z\setminus(X\cup D(X))})) + B_{X\cup D(X)} Q_{X\cup Sib(X)} - \log(\mathbb{P}(\chi^{Sib(X)})) \\
&= Q_{Z\setminus(X\cup D(X))} + B_{X\cup D(X)} Q_{X\cup Sib(X)} - Q_{Sib(X)}.
\end{aligned} \tag{K.3}$$

Now, substracting Equation (K.1) to Equation (K.3), and using Equation (H.1),

$$\begin{aligned}
Q_Z(\mathbb{1} - B_X) &= B_{X\cup D(X)} Q_{X\cup Sib(X)\cup D(X)\cup A(X)} - Q_{X\cup Sib(X)\cup D(X)\cup A(X)} \\
&= B_{Z\setminus(X\cup Sib(X)\cup D(X)\cup A(X))} Q_Z - B_{Z\setminus(Sib(X)\cup A(X))} Q_Z.
\end{aligned}$$

This leads to

$$Q_Z = B_{X\cup Sib(X)} Q_Z + B_{Z\setminus(X\cup Sib(X)\cup A(X)\cup D(X))} Q_Z - B_{Z\setminus(Sib(X)\cup A(X))} Q_Z = \beta_X Q_Z,$$

which ends the proof of Lemma K.1. ■

References.

[1] W. G. Bade, *On Boolean algebras of projections and algebras of operators*, Transactions of the American Mathematical Society, 80 (1955), pp. 345–345, https://doi.org/10.1090/S0002-9947-1955-0073954-0, http://www.ams.org/jourcgi/jour-getitem?pii=S0002-9947-1955-0073954-0 (accessed 2021-07-26).

[2] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, *Geometric deep learning: going beyond Euclidean data*, IEEE Signal Processing Magazine, 34 (2017), pp. 18–42, https://doi.org/10.1109/MSP.2017.2693418, http://arxiv.org/abs/1611.08097 (accessed 2021-06-07). arXiv: 1611.08097.

[3] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*, May 2017, http://arxiv.org/abs/1606.00915 (accessed 2022-09-15). arXiv:1606.00915 [cs].

[4] G. F. Cooper, *The computational complexity of probabilistic inference using bayesian belief networks*, Artificial Intelligence, 42 (1990), pp. 393–405, https://doi.org/10.1016/0004-3702(90)90060-D, https://www.sciencedirect.com/science/article/pii/000437029090060D (accessed 2022-08-29).

[5] I. Csiszár and Z. Talata, *Consistent estimation of the basic neighborhood of Markov random fields*, The Annals of Statistics, 34 (2006), pp. 123–145, https://doi.org/10.1214/009053605000000912, https://projecteuclid.org/journals/annals-of-statistics/volume-34/

issue-1/Consistent-estimation-of-the-basic-neighborhood-of-Markov-random-fields/10.1214/009053605000000912.full (accessed 2022-05-30). Publisher: Institute of Mathematical Statistics.

[6] J. Hammersley and P. Clifford, *Markov Fields On Finite Graphs And Lattices*. 1971, http://www.statslab.cam.ac.uk/~grg/books/hammfest/hamm-cliff.pdf (accessed 2020-12-17).

[7] M. Hochster, *Prime ideal structure in commutative rings*, Transactions of the American Mathematical Society, 142 (1969), pp. 43–60, https://doi.org/10.1090/S0002-9947-1969-0251026-X, https://www.ams.org/tran/1969-142-00/S0002-9947-1969-0251026-X/ (accessed 2022-09-16).

[8] Z. Huang, W. Xu, and K. Yu, *Bidirectional LSTM-CRF Models for Sequence Tagging*, Aug. 2015, https://doi.org/10.48550/arXiv.1508.01991, http://arxiv.org/abs/1508.01991 (accessed 2022-09-15). arXiv:1508.01991 [cs].

[9] E. V. Huntington, *Sets of Independent Postulates for the Algebra of Logic*, Transactions of the American Mathematical Society, 5 (1904), pp. 288–309, https://doi.org/10.2307/1986459, https://www.jstor.org/stable/1986459 (accessed 2022-09-12). Publisher: American Mathematical Society.

[10] B. Jonnson and A. Tarski, *Boolean Algebras with Operators*, American Journal of Mathematics, 74 (1952), p. 127, https://doi.org/10.2307/2372074, https://www.jstor.org/stable/2372074?origin=crossref (accessed 2022-09-16).

[11] T. N. Kipf and M. Welling, *Semi-Supervised Classification with Graph Convolutional Networks*, Feb. 2017, https://doi.org/10.48550/arXiv.1609.02907, http://arxiv.org/abs/1609.02907 (accessed 2022-08-26). arXiv:1609.02907 [cs, stat].

[12] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, *Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data*, in Proceedings of the Eighteenth International Conference on Machine Learning, ICML '01, San Francisco, CA, USA, June 2001, Morgan Kaufmann Publishers Inc., pp. 282–289.

[13] S. Li, L. Wu, S. Feng, F. Xu, F. Xu, and S. Zhong, *Graph-to-Tree Neural Networks for Learning Structured Input-Output Translation with Applications to Semantic Parsing and Math Word Problem*, Oct. 2020, https://doi.org/10.48550/arXiv.2004.13781, http://arxiv.org/abs/2004.13781 (accessed 2022-08-29). arXiv:2004.13781 [cs, stat].

[14] N. Peng, H. Poon, C. Quirk, K. Toutanova, and W.-t. Yih, *Cross-Sentence N-ary Relation Extraction with Graph LSTMs*, Transactions of the Association for Computational Linguistics, 5 (2017), pp. 101–115, https://doi.org/10.1162/tacl_a_00049, https://doi.org/10.1162/tacl_a_00049 (accessed 2022-09-15).

[15] H. A. Priestley, *Ordered Topological Spaces and the Representation of Distributive Lattices*, Proceedings of the London Mathematical Society, s3-24 (1972), pp. 507–530, https://doi.org/10.1112/plms/s3-24.3.507, http://doi.wiley.com/10.1112/plms/s3-24.3.507 (accessed 2022-09-16).

[16] W. Ricker, *Operator algebras generated by commuting projections: a vector measure approach*, no. 1711 in Lecture notes in mathematics, Springer, Berlin ; New York, 1999.

[17] D. Roth, *On the hardness of approximate reasoning*, Artificial Intelligence, 82 (1996), pp. 273–302, https://doi.org/10.1016/0004-3702(94)00092-1, https://linkinghub.elsevier.com/retrieve/pii/0004370294000921 (accessed 2022-08-29).

[18] J. H. Rubinstein, *Minimal prime ideals and compactifications*, Journal of the Australian Mathematical Society, 13 (1972), pp. 423–432, https://doi.org/10.1017/S1446788700009162, https://www.cambridge.org/core/product/identifier/S1446788700009162/type/journal_article (accessed 2022-09-16).

[19] M. H. Stone, *Subsumption of the Theory of Boolean Algebras under the Theory of Rings*, Proceedings of the National Academy of Sciences, 21 (1935), pp. 103–105, https://doi.org/10.1073/pnas.21.2.103, https://www.pnas.org/doi/10.1073/pnas.21.2.103 (accessed 2022-09-12). Publisher: Proceedings of the National Academy of Sciences.

[20] M. H. Stone, *The Theory of Representation for Boolean Algebras*, Transactions of the American Mathematical Society, 40 (1936), p. 37, https://doi.org/10.2307/1989664, https://www.jstor.org/stable/1989664?origin=crossref (accessed 2022-08-25).

[21] K. S. Tai, R. Socher, and C. D. Manning, *Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks*, May 2015, https://doi.org/10.48550/arXiv.1503.00075, http://arxiv.org/abs/1503.00075 (accessed 2022-08-29). arXiv:1503.00075 [cs].

[22] R. Talak, S. Hu, L. Peng, and L. Carlone, *Neural Trees for Learning on Graphs*, Oct. 2021, https://doi.org/10.48550/arXiv.2105.07264, http://arxiv.org/abs/2105.07264 (accessed 2022-08-29). arXiv:2105.07264 [cs].

[23] J. Topping, F. Di Giovanni, B. P. Chamberlain, X. Dong, and M. M. Bronstein, *Understanding over-squashing and bottlenecks on graphs via curvature*, Mar. 2022, https://doi.org/10.48550/arXiv.2111.14522, http://arxiv.org/abs/2111.14522 (accessed 2022-08-26). arXiv:2111.14522 [cs, stat].

[24] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, *Graph Attention Networks*, Feb. 2018, https://openreview.net/forum?id=rJXMpikCZ (accessed 2021-08-02).

[25] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, *How Powerful are Graph Neural Networks?*, Feb. 2019, http://arxiv.org/abs/1810.00826 (accessed 2022-08-26). arXiv:1810.00826 [cs, stat].

[26] M. Zhang and Y. Chen, *Link Prediction Based on Graph Neural Networks*, Nov. 2018, https://doi.org/10.48550/arXiv.1802.09691, http://arxiv.org/abs/1802.09691 (accessed 2022-09-15). arXiv:1802.09691 [cs, stat].

[27] M. Zitnik, M. Agrawal, and J. Leskovec, *Modeling polypharmacy side effects with graph convolutional networks*, Bioinformatics, 34 (2018), pp. i457–i466, https://doi.org/10.1093/bioinformatics/bty294, https://doi.org/10.1093/bioinformatics/bty294 (accessed 2022-09-15).