



# Self-adhesivity in lattices of abstract conditional independence models



Tobias Boege<sup>a,\*</sup>, Janneke H. Bolt<sup>b,c</sup>, Milan Studený<sup>d</sup>

<sup>a</sup> KTH Royal Institute of Technology, Department of Mathematics, 100 44, Stockholm, Sweden

<sup>b</sup> Utrecht University, Department of Information and Computing Sciences, 3584 CS Utrecht, The Netherlands

<sup>c</sup> Open University of the Netherlands, Faculty of Science, 6419 AT, Heerlen, The Netherlands

<sup>d</sup> Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 182 00, Prague, Czech Republic

## ARTICLE INFO

### Article history:

Received 18 March 2024

Received in revised form 2 October 2024

Accepted 7 October 2024

Available online xxxx

### Keywords:

Self-adhesivity

Conditional independence

Semi-graphoid

Lattice (order theory)

Pseudo-closed element

Boolean satisfiability

## ABSTRACT

We introduce an algebraic concept of the frame for abstract *conditional independence* (CI) models, together with basic operations with respect to which such a frame should be closed: copying and marginalization. Three standard examples of such frames are (discrete) probabilistic CI structures, semi-graphoids and structural semi-graphoids. We concentrate on those frames which are closed under the operation of set-theoretical intersection because, for these, the respective families of CI models are lattices. This allows one to apply the results from lattice theory and formal concept analysis to describe such families in terms of implications among CI statements.

The central concept of this paper is that of *self-adhesivity* defined in algebraic terms, which is a combinatorial reflection of the self-adhesivity concept studied earlier in context of polymatroids and information theory. The generalization also leads to a self-adhesivity operator defined on the meta-level of CI frames. We answer some of the questions related to this approach and raise other open questions.

The core of the paper is in computations. The combinatorial approach to computation might overcome some memory and space limitation of software packages based on polyhedral geometry, in particular, if SAT solvers are utilized. We characterize some basic CI families over 4 variables in terms of canonical implications among CI statements. We apply our method in information-theoretical context to the task of entropic region demarcation over 5 variables.

© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

This paper is devoted to the concept of *self-adhesivity* in the context of abstract *conditional independence* (CI) structures, which generalize/approximate classic probabilistic CI structures [35]. The concepts of adhesivity and self-adhesivity have been introduced by Matúš [29] in the context of *polymatroids*, which are special set functions playing a crucial role in combinatorial optimization [14]. Matúš's motivation came from information theory [41], more specifically, from the area of information-theoretical inequalities. The self-adhesivity concept for polymatroids is a kind of abstraction of the method used to derive the first non-Shannon information inequality [43], which method later became known in information-theoretical community as the so-called *copy lemma* [13]. The method led to the application of tools of polyhedral geometry (and linear programming solvers) to derive new inequalities [9].

\* Corresponding author.

E-mail addresses: [post@taboeger.de](mailto:post@taboeger.de) (T. Boege), [j.h.bolt@uu.nl](mailto:j.h.bolt@uu.nl) (J.H. Bolt), [studený@utia.cas.cz](mailto:studený@utia.cas.cz) (M. Studený).

In this paper we introduce the concept of self-adhesivity for abstract CI structures, which is a combinatorial reflection of the self-adhesivity concept for polymatroids. The idea comes from [3] and has already been applied in the context of Gaussian CI structures [4]. Nonetheless, our paper puts even more emphasis on the abstract view on (the families of) CI models and utilizes the results of *lattice theory* [2] for this purpose. More specifically, every finite lattice can alternatively be described in terms of a closure operation, or in terms of (abstract) functional dependence relation [24]. Thus, one can utilize the results from the theory of *formal concept analysis* [16]. This leads to an analogy with mathematical logic and to the effort to describe families of CI models in an axiomatic way, using implications among CI statements.

We introduce the notion of an abstract frame for CI structures and give several specific examples of such CI frames. We also introduce relevant algebraic operations with respect to which these abstract CI frames should be closed: these are copying and marginalization. From the point of view of our paper, the substantial algebraic operation is the set-theoretical *intersection*, which ensures that, for a fixed (finite) set of variables, the family of respective CI models forms a finite lattice. Then we introduce the concept of self-adhesivity, which is relative to a CI frame. The self-adhesivity concept induces special *self-adhesivity operator* of the meta-level of CI frames. This operator assigns a new CI frame to a starting CI frame and this new frame is a shrinkage of the original frame. In case the starting CI frame is an outer approximation of the classic probabilistic CI frame, the application of the self-adhesivity operator yields an even tighter outer approximation of the probabilistic CI frame.

This leads to a method to derive stronger valid probabilistic implications among CI statements on basis of previously known valid implications. The point of the story is that newly derived implications over  $n$  variables are obtained on basis of known implications over a higher number of variables than  $n$  through self-adhesivity.

The core of the paper are computations. Their goal was the description of several lattices of abstract CI models over a low number of variables. There are two methodologically different options: the lattices can

- either be described by listing meet-irreducible CI models,
- or in terms of implications among CI statements.

Our approach also allows one to derive such CI implications directly, by means of performing the computations on the combinatorial level of CI frames. To contrast with Matúš's original concept of self-adhesivity for polymatroids, note that the dimension of the polymatroidal cone grows exponentially in  $n$  and the number of CI statements grows even faster. Still, our computational methods based on SAT solvers can deal with this increase much better than the linear programming solvers used in the polymatroid setup.

In the paper, a specific example of the application of our combinatorial method is given. We applied the method to the question of identification of those extreme rays of the polymatroidal cone over 5 variables which are *not entropic*, which means that they do not have probabilistic representation. It appears that most of the 1319 permutational types of these extreme rays lie outside the entropic region. More specifically, it follows from our computations that at most 154 of types of such extreme rays may be entropic. The proof is entirely computational and takes about 330 s to execute on the author's laptop. Throughout the paper we also formulate remaining open mathematical questions related to our approach.

The structure of the paper is as follows. In Section 2, we define our concept of an abstract CI frame together with basic algebraic operations applied to CI models. We also recall those particular concepts and results from lattice theory on which our algorithms are based and which allow one to introduce frame-based closures of CI models. In Section 3 we present examples of such abstract CI frames. The concept of self-adhesion relative to such an abstract CI frame is introduced in Section 4, together with related basic observations and two other relevant algebraic operations applied to CI models. Section 5 is devoted to the self-adhesivity operator and the idea of its iterating. In Section 6 we review our tools for computations. The results of our computational experiments are presented in Section 7 and their application in information-theoretical context is explained in Section 8. Concluding remarks are presented in Section 9. More detailed information about our computations, source code and resulting data files can be found on the website: <https://cinet.link/data/selfadhe-lattices>.

## 2. Basic concepts

In this section we introduce both our notational conventions and terminology. Recall that **CI** is an abbreviation for *conditional independence*.

Given a set  $X$ , the symbol  $\mathcal{P}(X) := \{S : S \subseteq X\}$  will denote its *power set*, that is, the set of all subsets of  $X$ . The upper case letters like  $K, L, M, N, O$  will denote finite sets of *variables* of interest. In our intended probabilistic interpretation, their elements correspond to random variables. Additionally, two special notational conventions will be used. Given a variable  $i \in N$ , the symbol  $i$  will also denote the singleton set  $\{i\}$  and, given two variable sets  $K$  and  $L$ , the juxtaposition  $KL$  will be a shortened notation for their union  $K \cup L$ . Given a bijection  $\varphi : N \rightarrow M$  and  $L \subseteq N$  the restriction of  $\varphi$  to  $L$  will be denoted by  $\varphi|_L$ , the identical bijection of  $L$  onto itself by  $\text{id}_L$ .

### 2.1. Abstract families of CI models and the probabilistic CI frame

An elementary *CI statement* over a finite variable set  $N$  is the statement claiming that, for some  $K \subseteq N$  and distinct  $i, j \in N \setminus K$ , the variable  $i$  is conditionally independent of the variable  $j$  given  $K$ , traditionally denoted by  $i \perp\!\!\!\perp j \mid K$ . A converse conditional *dependence statement* (= negation) is the claim that  $i$  and  $j$  are conditionally dependent given  $K$ , traditionally denoted by  $i \not\perp\!\!\!\perp j \mid K$ . Thus, a triplet  $(i, j|K)$  of pairwise disjoint subsets of  $N$  can be interpreted either as an independence statement or as a dependence statement.

Because, in this paper, we do not deal with “mixed” CI structures, involving both dependence and independence statements, we accept a convention that our sketchy notation  $ij|K$  for a triplet  $(i, j|K)$  will be interpreted as the CI statement  $i \perp\!\!\!\perp j \mid K$ . Moreover, in our paper we consider only *symmetric* CI concepts in the sense that  $i \perp\!\!\!\perp j \mid K$  is equivalent to  $j \perp\!\!\!\perp i \mid K$ . Every such a “symmetric” CI statement can mathematically be described by an ordered pair  $L \mid K$  of sets  $L, K \subseteq N$ , where  $L \cap K = \emptyset$  and  $|L| = 2$ ; specifically we have  $L = \{i, j\}$ . So, our second convention is that  $ij|K$  means the same as  $ji|K$ .

Note that this way of introducing CI statements implicitly means that any CI statement over a variable set  $N$  is automatically also a CI statement over any variable set  $M$  with  $N \subseteq M$ .

**Definition 2.1** (*Abstract CI Statements, Models, Families, Frames*).

Given a finite variable set  $N$ , the symbol  $\kappa(N)$  will denote the collection of all ordered pairs  $L \mid K$ , where  $L, K \subseteq N$ ,  $L \cap K = \emptyset$ , and  $|L| = 2$ .

- The elements of  $\kappa(N)$  will be interpreted as (CI) *statements* over  $N$ . They will usually be denoted by explicit writing:  $ij|K \in \kappa(N)$ .
- A subset of  $\kappa(N)$  will be interpreted as a (CI) *model* over  $N$ . Models will be denoted using the calligraphic font:  $\mathcal{M} \subseteq \kappa(N)$  or  $\mathcal{M} \in \mathcal{P}(\kappa(N))$ .
- A class of (CI) models over  $N$  involving the (largest possible) model  $\kappa(N)$  will be a (CI) *family* over  $N$ . Thus, a family is a subset of the power set of  $\kappa(N)$ . Families will be denoted using the fractal font and will have the variable sets as their arguments:  $\mathfrak{f}(N) \in \mathcal{P}(\kappa(N))$ .

By a (CI) *frame* we will understand a mapping which assigns a (CI) family over  $N$  to each finite (variable) set  $N$ . Frames will be denoted using the fractal font without an argument: mapping  $\mathfrak{f} : N \mapsto \mathfrak{f}(N) \in \mathcal{P}(\mathcal{P}(\kappa(N)))$ . Given two CI frames  $\mathfrak{f}_1$  and  $\mathfrak{f}_2$ , we write  $\mathfrak{f}_1 \subseteq \mathfrak{f}_2$  to mean  $\mathfrak{f}_1(N) \subseteq \mathfrak{f}_2(N)$  for any (finite) variable set  $N$ .

The reader can recognize a clear hierarchy: a (CI) statement is an element of  $\kappa(N)$ , a model is an element of the power set of  $\kappa(N)$  and a family is an element of the power set of the power set of  $\kappa(N)$ .<sup>1</sup>

Note that  $\kappa(N) = \emptyset$  if  $N$  is empty or a singleton. In particular, only one (CI) model over  $N$  exists if  $|N| \leq 1$ , namely the empty set/model. Moreover, only one (CI) family over  $N$  exists if  $|N| \leq 1$  because of our explicit requirement that  $\mathfrak{f}(N) \in \mathcal{P}(\kappa(N))$  in the definition of a family.

The definition of a CI model is relative to a variable set  $N$  and, strictly speaking, the model is a pair  $(N, \mathcal{M})$  with  $\mathcal{M} \subseteq \kappa(N)$ . For example, the empty set might be considered as a CI model over  $\{a, b\}$  or over  $\{a, b, c\}$ . To simplify notation, we will usually rely on context to identify the variable set or mention it explicitly by means of the “frame notation”  $\mathcal{M} \in \mathfrak{f}(N)$ . If two variable sets are in inclusion  $N \subseteq M$ , then any CI model  $\mathcal{M}$  over  $N$  may be regarded as a model over  $M$  as well. In the context of a frame, this ability to transition to larger variable sets is formalized by the operation of *ascetic extension* defined below.

**Example 2.2.** Put  $N := \{a, b, c, d\}$  and  $\mathcal{M} := \{ab|\emptyset, bc|\emptyset\} \subseteq \kappa(N)$ . Then  $N^{\mathcal{M}} = \{a, b, c\}$  and  $\mathcal{M}$  is a CI model over this strict subset of  $N$ . Nonetheless,  $\mathcal{M}$  is not a CI model over  $\{a, b\}$  nor over  $\{b, c, e\}$ , where  $e$  is a variable outside  $N$ . On the other hand,  $\mathcal{M}$  is a model over any variable set  $M$  with  $\{a, b, c\} \subseteq M$ .

We now give a prototypical example of an abstract CI frame, which is the frame of discrete probabilistic CI structures. A *discrete random vector* over a variable set  $N$  is an indexed collection  $\xi = [\xi_i]_{i \in N}$  of random variables, each  $\xi_i$  taking its values in an individual finite sample space  $X_i$ , which is necessarily non-empty then. The joint sample space for  $\xi$  is  $X_N := \prod_{i \in N} X_i$ . Given  $\xi$  and  $I \subseteq N$ , the respective *marginal density* of  $\xi$  can be defined as function  $p_I : X_N \rightarrow [0, 1]$  such that the value  $p_I(x)$  at  $x = [x_i]_{i \in N} \in X_N$  is the probability of having  $\xi_i = x_i$  for each  $i \in I$ . Given  $ij|K \in \kappa(N)$ , we have  $i \perp\!\!\!\perp j \mid K \mid \xi$  if

$$\forall x \in X_N \quad p_{ijK}(x) \cdot p_K(x) = p_{iK}(x) \cdot p_{jK}(x).$$

The CI model induced by  $\xi$  is the set  $\mathcal{M}_\xi := \{ij|K \in \kappa(N) : i \perp\!\!\!\perp j \mid K \mid \xi\}$ . The family of *discrete probabilistic models* over  $N$  is then

$$\text{pr}(N) := \{\mathcal{M}_\xi : \xi \text{ is a discrete random vector over } N\}.$$

<sup>1</sup> A frame corresponds to what was called a “property” in the PhD thesis [3, §2.2].

Observe that  $\kappa(N)$  belongs to  $\text{pr}(N)$  because it is induced by a random vector with completely independent components. This defines the *discrete probabilistic CI frame*  $\text{pr}$ .

It will be convenient to deal with CI statements with more variables in place of the singletons  $i$  and  $j$ . Although the treatment of (probabilistic) CI statements can be extended to this more general case, in this paper the notation  $I \perp\!\!\!\perp J \mid K$  for disjoint subsets of  $N$  will be considered only as an abbreviation.

**Definition 2.3** (*Global CI Statement*).

Let  $N$  be a variable set and  $I, J, K \subseteq N$  pairwise disjoint subsets of it. Given a CI model  $\mathcal{M} \subseteq \kappa(N)$  over  $N$ , we will write  $I \perp\!\!\!\perp J \mid K [\mathcal{M}]$  if

$$\forall i \in I \quad \forall j \in J \quad \forall L : K \subseteq L \subseteq IJK \setminus ij \quad ij|L \in \mathcal{M}. \tag{1}$$

We interpret this as a global CI statement encoded within the CI model  $\mathcal{M}$ . In our examples and catalogues we also use a shorthand  $[I, J|K]$  to denote the list of (elementary) CI statements described in (1).

Note that, in the context of the probabilistic frame  $\text{pr}$ , the abbreviating condition  $I \perp\!\!\!\perp J \mid K [\mathcal{M}_\xi]$  is equivalent to the requirement that  $p_{IJK}(x) \cdot p_K(x) = p_{IK}(x) \cdot p_{JK}(x)$  holds for every  $x \in X_N$ , which requirement is interpreted as the (global) probabilistic CI statement  $I \perp\!\!\!\perp J \mid K [\xi]$ ; see [35, Lemma 2.2].

The next step is to introduce five algebraic operations on abstract CI models. The goal is to study various CI frames which are closed under such operations.

**Definition 2.4** (*Copying, Marginalization, Intersection, Duality, Ascetic Extension*).

- Any bijection  $\varphi : N \rightarrow M$  between two variable sets  $N$  and  $M$  can naturally be extended to a bijection between their power sets and to a bijection between  $\kappa(N)$  and  $\kappa(M)$ : put  $\varphi(ij|K) := \varphi(ij)|\varphi(K)$  for every  $ij|K \in \kappa(N)$ .
- Any model  $\mathcal{M} \subseteq \kappa(N)$  over  $N$  is then transformed by  $\varphi$  to a model over  $M$ : put

$$\varphi(\mathcal{M}) := \{ \varphi(ij|K) : ij|K \in \mathcal{M} \} \subseteq \kappa(M).$$

This is the  $\varphi$ -copy of  $\mathcal{M}$ . We say that a CI frame  $\mathfrak{f}$  is closed under *copying* if, for every model  $\mathcal{M} \in \mathfrak{f}(N)$  and every bijection  $\varphi : N \rightarrow M$ , one has  $\varphi(\mathcal{M}) \in \mathfrak{f}(M)$ .

- Given a model  $\mathcal{M} \subseteq \kappa(N)$  over  $N$  and  $M \subseteq N$ , its *marginal* for  $M$ , denoted by  $\mathcal{M}^{\downarrow M}$ , is simply its restriction to  $\kappa(M)$ : put  $\mathcal{M}^{\downarrow M} := \mathcal{M} \cap \kappa(M)$ . A CI frame  $\mathfrak{f}$  is closed under *marginalization* if, for every model  $\mathcal{M} \in \mathfrak{f}(N)$  and every subset  $M \subseteq N$ , one has  $\mathcal{M}^{\downarrow M} \in \mathfrak{f}(M)$ .
- Given two models  $\mathcal{M}, \mathcal{R} \subseteq \kappa(N)$  over the same set  $N$ , their (set-theoretical) *intersection* is the model  $\mathcal{M} \cap \mathcal{R}$ , again a model over  $N$ . A CI frame  $\mathfrak{f}$  will be called closed under *intersection* if, for every two models  $\mathcal{M}, \mathcal{R} \in \mathfrak{f}(N)$ , one has  $\mathcal{M} \cap \mathcal{R} \in \mathfrak{f}(N)$ .
- Given a variable set  $N$ , the duality map from  $\kappa(N)$  to  $\kappa(N)$  is as follows:  $ij|K \in \kappa(N)$  is assigned the pair  $(ij|K)^* := ij|L$ , where  $L := N \setminus (K \cup ij)$ . Given a model  $\mathcal{M} \subseteq \kappa(N)$ , the *dual model*  $\mathcal{M}^*$  is its image by this mapping. A CI frame  $\mathfrak{f}$  is closed under *duality* if  $\mathcal{M} \in \mathfrak{f}(N)$  implies  $\mathcal{M}^* \in \mathfrak{f}(N)$ .
- Given a model  $\mathcal{M} \subseteq \kappa(N)$  over  $N$  and  $N \subseteq M$ , an *extension* of  $\mathcal{M}$  to  $M$  is any model  $\mathcal{R} \subseteq \kappa(M)$  over  $M$  such that  $\mathcal{R}^{\downarrow N} = \mathcal{M}$ . A CI frame  $\mathfrak{f}$  is closed under *extension* if, for every model  $\mathcal{M} \in \mathfrak{f}(N)$  and every  $M$  with  $M \supseteq N$ , there exists at least one extension  $\mathcal{R}$  of  $\mathcal{M}$  to  $M$  with  $\mathcal{R} \in \mathfrak{f}(M)$ .
- Given  $\mathcal{M} \subseteq \kappa(N)$  and  $N \subseteq M$ , an extension  $\mathcal{R} \subseteq \kappa(M)$  of  $\mathcal{M}$  to  $M$  will be called *ascetic* if  $\mathcal{R} \setminus \mathcal{M} = \emptyset$ . A CI frame  $\mathfrak{f}$  is closed under *ascetic extension* if, for every model  $\mathcal{M} \in \mathfrak{f}(N)$  and every variable set  $M$  with  $M \supseteq N$ , one has  $\mathcal{M} \in \mathfrak{f}(M)$ , that is, the ascetic extension of  $\mathcal{M}$  to  $M$  belongs to  $\mathfrak{f}(M)$ .

Since  $\kappa(N) = \emptyset$  if  $|N| \leq 1$ , the concepts of copying and duality make real sense only in case  $|N| \geq 2$ . On the other hand, to introduce self-adhesion in Section 4 one needs marginalization to sets  $M$  with  $|M| \leq 1$ .

Note that if  $N \subseteq M$  then every model  $\mathcal{M}$  over  $N$  has an ascetic extension to  $M$ , namely  $\mathcal{M}$  itself, viewed as a CI model over  $M$ . Clearly, this is the least possible extension of  $\mathcal{M}$  to  $M$  and, thus, unique. The word ‘‘ascetic’’ was chosen because, in comparison with other extending algebraic operations, it adds nothing to  $\mathcal{M}$ . Clearly, a CI frame  $\mathfrak{f}$  is closed under ascetic extension iff  $N \subseteq M$  implies  $\mathfrak{f}(N) \subseteq \mathfrak{f}(M)$ . We leave it to the reader as an easy exercise to show that, if a CI frame  $\mathfrak{f}$  is closed under extension and intersection, then it is closed under ascetic extension iff  $N \subseteq M$  implies  $\kappa(N) \in \mathfrak{f}(M)$ .

Two basic algebraic operations with abstract CI models are copying and marginalization. Most of the considered CI frames in this paper are also closed under intersection, which implies that the respective families of models are lattices and one can introduce frame-based closures for CI models (see Section 2.3).

Our prototypical example of a CI frame is closed under all but one algebraic operations considered.

**Lemma 2.5.** *The frame  $\text{pr}$  is closed under copying, marginalization, intersection, and ascetic extension.*

**Proof.** Given a discrete random vector  $\xi = [\xi_i]_{i \in N}$  and a bijection  $\varphi : N \rightarrow M$ , the random vector  $\eta := [\xi_{\varphi^{-1}(j)}]_{j \in M}$  induces the  $\varphi$ -copy of  $\mathcal{M}_\xi$ : this is because  $i \perp\!\!\!\perp j \mid K [\xi]$  if and only if  $\varphi(i) \perp\!\!\!\perp \varphi(j) \mid \varphi(K) [\eta]$ . Given  $\xi = [\xi_i]_{i \in N}$  and  $M \subseteq N$ , the random sub-vector  $\xi_M := [\xi_i]_{i \in M}$  induces the marginal model  $(\mathcal{M}_\xi)^{\downarrow M}$ .

Let  $\xi^1 = [\xi_i^1]_{i \in N}$  and  $\xi^2 = [\xi_i^2]_{i \in N}$  be discrete random vectors over  $N$ . Take a bijection  $\varphi : N \rightarrow M$  with  $N \cap M = \emptyset$  and construct an extended random vector  $\xi = [\xi_i]_{i \in NM}$  with stochastically independent sub-vectors  $\xi_N := \xi^1 = [\xi_i^1]_{i \in N}$  and  $\xi_M := [\xi_{\varphi^{-1}(j)}^2]_{j \in M}$ . Define a random vector  $\eta = [\eta_i]_{i \in N}$  over  $N$  where each  $\eta_i := [\xi_i, \xi_{\varphi(i)}]$  is a composite random variable. The construction allows one to observe that, for any  $ij|K \in \mathcal{K}(N)$ , one has  $i \perp\!\!\!\perp j | K [\eta]$  if and only if both  $i \perp\!\!\!\perp j | K [\xi^1]$  and  $i \perp\!\!\!\perp j | K [\xi^2]$ . This gives  $\mathcal{M}_\eta = \mathcal{M}_{\xi^1} \cap \mathcal{M}_{\xi^2}$ .

To show that  $\text{pr}$  is closed under extension consider  $N \subseteq M$  and a discrete random vector  $\xi = [\xi_i]_{i \in N}$  and extend it to a random vector  $\eta = [\eta_i]_{i \in M}$  with stochastically independent sub-vectors  $\eta_N := \xi$  and  $\eta_i, i \in M \setminus N$ . Because we already know that  $\text{pr}$  is closed under intersection, to show that it is closed under ascetic extension it remains to show that, if  $N \subseteq M$ , then  $\mathcal{K}(N) \in \text{pr}(M)$ . This leads to a construction of a random vector  $\eta = [\eta_i]_{i \in M}$  with stochastically independent components  $\eta_i, i \in N$ , while there is no other CI relation within  $\eta$ . A specific construction of such a random vector is left to the reader.  $\square$

On the other hand, the CI frame  $\text{pr}$  is not closed under duality. Recall a counter-example by Matúš from [26, end of §5].

**Example 2.6.** Given  $N := \{a, b, c, d\}$  and  $\mathcal{M} := \{ab|cd, ab|d, ab|c, cd|\emptyset\}$ , one has  $\mathcal{M} \in \text{pr}(N)$  and  $\mathcal{M}^* = \{ab|\emptyset, ab|c, ab|d, cd|ab\} \notin \text{pr}(N)$ . Indeed, a binary random vector uniformly distributed on four special configurations

$$(0, 0, 0, 0) \quad (0, 1, 0, 1) \quad (0, 1, 1, 0) \quad (1, 1, 1, 1)$$

of values for  $(x_a, x_b, x_c, x_d)$  implies the former fact. The latter fact follows from the property (I:13) in [37, §V.B] saying that the implication

$$[a \perp\!\!\!\perp b | \emptyset \ \& \ a \perp\!\!\!\perp b | c \ \& \ a \perp\!\!\!\perp b | d \ \& \ c \perp\!\!\!\perp d | ab] \Rightarrow a \perp\!\!\!\perp b | cd$$

holds in the context of discrete probabilistic CI structures.

**Remark.** This is to elucidate possible alternative ways to formalize CI models. The classic way of understanding probabilistic CI structures from [32], followed in [35], was to define an abstract CI model over  $N$  as the collection of global CI statements over  $N$  valid with respect to a random vector  $\xi$  over  $N$ . This approach leads to an alternative global mode, where  $\mathcal{K}(N)$  is replaced by the collection  $\mathcal{T}(N)$  of all triplets  $\langle I, J|K \rangle$  of pairwise disjoint subsets  $I, J, K$  of  $N$  and (global) CI models are subsets of  $\mathcal{T}(N)$ . Such an alternative approach is equivalent to the one discussed in our paper through the relation from Definition 2.3.

Non-equivalent descriptions of probabilistic CI structures are those that additionally deal with *functional dependencies among random variables*, as the one from [30]. Given a discrete random vector  $\xi$  over  $N, K \subseteq N$ , and  $i \in N \setminus K$ , the respective functional dependence statement is the claim that  $\xi_i$  is distribution-equivalent to a function of  $\xi_K := [\xi_j]_{j \in K}$ . The point is that any such a functional dependence statement within  $\xi$  can be interpreted as a generalized CI statement  $i \perp\!\!\!\perp i | K [\xi]$ , namely that  $i$  is conditionally independent of itself given  $K$ . The way to formalize such wider CI structures is to replace  $\mathcal{K}(N)$  by the collection  $\mathcal{K}^\bullet(N)$  of ordered pairs  $L | K$ , where  $L, K \subseteq N, L \cap K = \emptyset$ , and  $1 \leq |L| \leq 2$ . Any pair  $i | K$  then corresponds to a functional dependence statement claiming that  $i \in N \setminus K$  is “determined” by  $K \subseteq N$ . Subsets of  $\mathcal{K}^\bullet(N)$  were called *augmented* CI models over  $N$  [37, §II.B], represented in the local/elementary mode. An equivalent version of this would be to replace  $\mathcal{K}^\bullet(N)$  by the collection  $\mathcal{T}^\bullet(N)$  of all triplets  $\langle I, J|K \rangle$  of subsets  $I, J, K$  of  $N$ . The subsets of  $\mathcal{T}^\bullet(N)$  could be viewed as augmented CI models, represented in the global mode. However, we do not consider augmented CI models with functional dependence statements in this paper.

## 2.2. Tutorial on galois connections and formal concept analysis

This part contains an overview of particular results from lattice theory and formal concept analysis we later apply in our context of abstract CI frames (see Section 2.3) to establish frame-based implications of CI models. We have included such a summary in our paper because a common reader need not be familiar with these finer results, perhaps well-known in lattice-theoretical community, but not in general. The reason is that the consistency of our algorithms (Appendix A) follows just from these facts.

The main message of this tutorial to be passed towards the reader is that any finite lattice, possibly with a huge number of elements, can (efficiently) be characterized in two methodologically different ways:

- in terms of irreducible elements (= an inner description in the sense that the characterization only depends on the lattice itself),
- in terms of implicational generators (= an outer description defined in case the lattice is specifically embedded into a Boolean lattice).

Below we describe both approaches and recall relevant concepts from lattice theory [2] for this purpose.

Throughout this tutorial,  $X$  will be a general set. Facts from lattice theory hold for any such a set  $X$ . Finer concepts, which concern the implications between subsets of  $X$ , and results from formal concept analysis implicitly assume that  $X$  is finite. We are going to apply those observations in Section 2.3 in the case  $X := \mathcal{K}(N)$  for some finite variable set  $N$ .



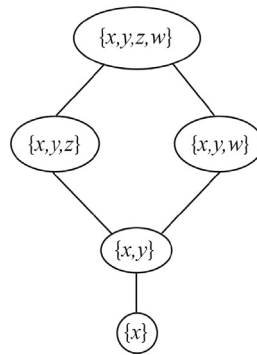


Fig. 1. A Moore family  $\{ \{x, y, z, w\}, \{x, y, z\}, \{x, y, w\}, \{x, y\}, \{x\} \}$ .

Given a set  $X$ , a *Moore family* of subsets of  $X$ , alternatively called a *closure system* (for  $X$ ), is a collection  $\mathcal{F} \subseteq \mathcal{P}(X)$  of subsets of  $X$  which is closed under (arbitrary) set intersection and includes the set  $X$  itself:

$$\mathcal{D} \subseteq \mathcal{F} \Rightarrow \bigcap \mathcal{D} \in \mathcal{F} \quad \text{with a convention that } \bigcap \mathcal{D} = X \text{ for } \mathcal{D} = \emptyset.$$

An elementary observation is that any Moore family  $(\mathcal{F}, \subseteq)$ , equipped with the set inclusion relation, is a complete lattice [2, Theorem 2 in §V.1] in which the meet operation is intersection.<sup>2</sup> Hence, the following simple facts can be derived.

**Observation 2.7.** *If a CI frame  $\mathfrak{f}$  is closed under intersection then, for each finite  $N$ , the family  $(\mathfrak{f}(N), \subseteq)$  is a finite lattice in which the meet operation is intersection. If  $\mathfrak{f}$  is also closed under copying and  $|N| = |M|$  then  $(\mathfrak{f}(N), \subseteq)$  and  $(\mathfrak{f}(M), \subseteq)$  are order-isomorphic and permutations of  $N$  yield automorphisms of  $(\mathfrak{f}(N), \subseteq)$ .*

**Proof.** By Definition 2.4, as  $\mathfrak{f}(N)$  is a Moore family, this is evident from what was said before.  $\square$

Let  $(\mathcal{L}, \leq)$  be a (finite) lattice with the greatest element denoted by  $\mathbf{1}$  and the meet/infimum operator denoted by  $\text{inf}$ . Let  $<$  denote the strict order relation. Given  $\lambda, \tau \in \mathcal{L}$ , one says that  $\tau$  covers  $\lambda$  and writes  $\lambda < \tau$  if  $\lambda < \tau$  and there is no  $\kappa \in \mathcal{L}$  with  $\lambda < \kappa < \tau$ . A *coatom* of  $\mathcal{L}$  is a sub-maximal element in  $\mathcal{L}$ , that is,  $\tau \in \mathcal{L}$  satisfying  $\tau < \mathbf{1}$ . An element  $\lambda \in \mathcal{L}$  is called (meet) *irreducible* if  $\lambda \neq \text{inf} \{ \tau \in \mathcal{L} : \lambda < \tau \}$ .<sup>3</sup> A well-known fact is that the set  $\mathcal{I}$  of (meet) irreducible elements in a finite lattice  $(\mathcal{L}, \leq)$  is the least set determining all elements of  $\mathcal{L}$  by means of the infimum operation in the following sense:

$$\forall \kappa \in \mathcal{L} \quad \exists \mathcal{S} \subseteq \mathcal{I} \quad \text{such that } \kappa = \text{inf} \{ \tau : \tau \in \mathcal{S} \}.$$

The coatoms are standard examples of meet irreducible elements in  $\mathcal{L}$ . In general, a lattice can have other irreducible elements. A lattice  $(\mathcal{L}, \leq)$  is called *coatomistic* if every element of  $\mathcal{L}$  is the infimum of a set of coatoms, which is another way of saying that there is no irreducible element in  $\mathcal{L}$  except the coatoms.

**Example 2.8.** Put  $X := \{x, y, z, w\}$  and consider a Moore family

$$\mathcal{F} := \{ \{x, y, z, w\}, \{x, y, z\}, \{x, y, w\}, \{x, y\}, \{x\} \},$$

of subsets of  $X$ , depicted in Fig. 1 in the form of its Hasse diagram. The elements  $\{x, y, z\}$ ,  $\{x, y, w\}$  and  $\{x\}$  of the lattice  $(\mathcal{F}, \subseteq)$  are then irreducible, but only the elements  $\{x, y, z\}$  and  $\{x, y, w\}$  are its coatoms. Thus,  $(\mathcal{F}, \subseteq)$  is a finite lattice which is not coatomistic.

The first way to characterize a finite lattice  $(\mathcal{F}, \subseteq)$ , where  $\mathcal{F}$  is a Moore family of subsets of  $X$ , is to give the complete list of irreducible elements in  $\mathcal{F}$ . In our considered case they fall into equivalence classes:

**Observation 2.9.** *If a CI frame  $\mathfrak{f}$  is closed under copying and intersection then, for each finite variable set  $N$ , the meet irreducible sets in  $(\mathfrak{f}(N), \subseteq)$  (= irreducible models) fall into permutational equivalence classes (= relative to permutations of  $N$ ).*

**Proof.** By Observation 2.7, any permutation  $\pi : N \rightarrow N$  defines an order automorphism of  $(\mathfrak{f}(N), \subseteq)$ , the natural extension of  $\pi$  to  $\mathcal{K}(N)$ . Such a transformation preserves the irreducibility of an element of  $\mathfrak{f}(N)$ .  $\square$

<sup>2</sup> This is a universal construction of a complete lattice in the sense that any such a lattice is isomorphic to a Moore family and any finite lattice is isomorphic to a finite Moore family lattice. These conclusions can be derived from observations in [2, §V.9].

<sup>3</sup> Note that  $\mathbf{1}$  is not irreducible by this definition: it is the infimum of the empty set.

**Example 2.10** (Lattice of Probabilistic Models Over 3 Variables).

Let us take  $N := \{a, b, c\}$ ,  $X := \mathcal{K}(N)$ , and put  $\mathcal{F} := \text{pr}(N)$ . Either a human-made analysis or simple computations allows one to observe that the lattice  $(\text{pr}(N), \subseteq)$  has 22 elements in this case and they break into 10 permutational equivalence classes; for the result see [35, Figure 5.6]. There are 5 (meet) irreducible elements in  $\text{pr}(\{a, b, c\})$  which fall down into 3 permutational types:

- $\{ab|\emptyset, ac|\emptyset, bc|\emptyset\}$ ,
- $\{ab|c, ac|b, bc|a\}$ ,
- $\{ab|\emptyset, ab|c, ac|\emptyset, ac|b\}$  (with 2 additional permutational copies).

All these elements are coatoms; in particular, the lattice is coatomistic.

A kind of dual description of a complete lattice is based on the following concept. A *closure operation* on subsets of  $X$  is a mapping  $\text{cl} : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$  which is

- *extensive*:  $Y \subseteq \text{cl}(Y)$  for  $Y \subseteq X$ ,
- *isotone*:  $Y \subseteq Z \subseteq X \Rightarrow \text{cl}(Y) \subseteq \text{cl}(Z)$ , and
- *idempotent*:  $\text{cl}(\text{cl}(Y)) = \text{cl}(Y)$  for  $Y \subseteq X$ .

A set  $Y \subseteq X$  is then called *closed with respect to cl* if  $Y = \text{cl}(Y)$ . A basic facts are that the collection of subsets of  $X$  closed with respect to  $\text{cl}$  is a Moore family and any Moore family can be defined in this way [2, Theorem 1 in §V.1]. Specifically, given a Moore family  $\mathcal{F}$  of subsets of  $X$ , the formula

$$\text{cl}_{\mathcal{F}}(Y) := \bigcap \{Z \subseteq X : Y \subseteq Z \in \mathcal{F}\} \text{ for } Y \subseteq X,$$

defines a closure operation on subsets of  $X$  having  $\mathcal{F}$  as the collection of closed sets with respect to  $\text{cl}_{\mathcal{F}}$ ; see [16, Theorem 1 in §0.3] for the proof of this simple observation. Thus, for any  $Y \subseteq X$ , the closure  $\text{cl}_{\mathcal{F}}(Y)$  of  $Y$  relative to  $\mathcal{F}$  is the smallest element of  $\mathcal{F}$  containing  $Y$ .

**Example 2.11.** Consider the Moore family  $\mathcal{F}$  from Example 2.8. Then, for instance,  $\text{cl}_{\mathcal{F}}(\emptyset) = \{x\} = \text{cl}_{\mathcal{F}}(\{x\})$ ,  $\text{cl}_{\mathcal{F}}(\{y\}) = \{x, y\}$ , and  $\text{cl}_{\mathcal{F}}(\{x, z\}) = \{x, y, z\}$ .

The closure operation can also be interpreted as an *abstract functional dependence relation* on subsets of  $X$ , as explained in [24]. This is a binary relation  $\rightarrow$  on  $\mathcal{P}(X)$  satisfying three axioms (for  $Y, Z, W \subseteq X$ ):

- *reflexivity*:  $Z \subseteq Y \Rightarrow Y \rightarrow Z$ ,
- *transitivity*:  $[Y \rightarrow Z \ \& \ Z \rightarrow W] \Rightarrow Y \rightarrow W$ ,
- *union*:  $[Y \rightarrow Z \ \& \ Y \rightarrow W] \Rightarrow Y \rightarrow Z \cup W$ .

In case of finite  $X$ , these axioms are together equivalent to the well-known *Armstrong axioms* for functional dependence relations in database theory. Given a closure operation  $\text{cl}$  on subsets of a finite set  $X$ , define  $Y \rightarrow Z$  as  $Z \subseteq \text{cl}(Y)$  and observe it is an abstract functional dependence relation on subsets of  $X$ . Given a functional dependence relation  $\rightarrow$ , put  $\text{cl}(Y) := \bigcup \{Z \subseteq X : Y \rightarrow Z\}$  for  $Y \subseteq X$  and observe this defines a closure operation on subsets of  $X$  inducing the original functional dependence relation  $\rightarrow$ .

**Example 2.12.** Let us consider again the Moore family  $\mathcal{F}$  from Example 2.8 and its induced closure operation  $Y \mapsto \text{cl}_{\mathcal{F}}(Y)$  for  $Y \subseteq X$ . The respective functional dependence relation is quite huge. It certainly contains 16 implications saying that subsets of  $X$  imply their closures:

$$\begin{aligned} \emptyset \rightarrow \{x\}, \quad \{x\} \rightarrow \{x\}, \quad \{y\} \rightarrow \{x, y\}, \quad \{z\} \rightarrow \{x, y, z\}, \quad \{w\} \rightarrow \{x, y, w\}, \\ \dots \quad \{y, z, w\} \rightarrow \{x, y, z, w\}, \quad \{x, y, z, w\} \rightarrow \{x, y, z, w\}. \end{aligned}$$

Besides these, it also contains their natural consequences because of inclusion: for instance, it contains the implication  $\{y, z, w\} \rightarrow \{x, w\}$  which follows from the inclusion  $\{x, w\} \subseteq \text{cl}_{\mathcal{F}}(\{y, z, w\})$  (use reflexivity and transitivity for this).

**Remark.** To avoid possible terminological misunderstanding on the side of the reader note that functional dependencies within a random vector  $\xi$  mentioned in the remark concluding Section 2.1 can be viewed as a part of an abstract *probabilistic functional dependence relation*. Specifically, given a random vector  $\xi$  over a finite variable set  $N$ , the collection of pairs  $K \rightarrow I$ , where  $K, I \subseteq N$ ,  $K \cap I = \emptyset$ , and  $I \perp\!\!\!\perp K \mid \xi$  (augmented global CI statements) is an abstract functional dependence relation on subsets of  $X := N$ . One of the results in [24, Remark 4] is that every abstract functional dependence relation on subsets of a finite  $N$  is probabilistic in the above sense. So, these notions are indeed related.

Nonetheless, in our paper we ignore probabilistic functional dependencies within  $N$ . Instead, we deal with functional dependence on a meta-level of CI statements/models over  $N$ : we consider abstract functional dependence relations on subsets of  $X := \mathcal{K}(N)$ , where  $N$  is a finite variable set. To emphasize the different context, we also talk about “implications” between subsets of  $X$ , which are CI models over  $N$ .

The second way to characterize a finite lattice  $(\mathcal{F}, \subseteq)$ , where  $\mathcal{F}$  is a Moore family of subsets of  $X$ , is to give a suitable generator of the respective functional dependence relation  $\rightarrow$ . Before explaining the details we wish the reader to realize/observe that there is a straightforward relationship (= correspondence) between the functional dependence relations  $\rightarrow$  and Moore families  $\mathcal{F}$ .<sup>4</sup>

Specifically, any element  $(Y, Z)$  of  $\mathcal{P}(X) \times \mathcal{P}(X)$ , where  $X$  is a finite set, can be interpreted as a formal implication  $Y \rightarrow Z$ , meaning that the set  $Y$  of *antecedents* implies (all elements of) the set  $Z$  of *consequents*. One can introduce a particular binary relation  $\odot$  between such pairs  $(Y, Z)$  and subsets  $F$  of  $X$ : we say that  $Y \rightarrow Z$  is valid for  $F \subseteq X$  if  $Y \subseteq F$  implies  $Z \subseteq F$ . More technically:

$$\begin{aligned} \forall (Y \rightarrow Z) \in \mathcal{P}(X) \times \mathcal{P}(X) \quad \forall F \in \mathcal{P}(X) \\ (Y \rightarrow Z) \odot F \quad := \quad [ Y \setminus F \neq \emptyset \text{ or } Z \subseteq F ]. \end{aligned}$$

Next we introduce the so-called *Galois connections* between subsets of  $\mathcal{P}(X) \times \mathcal{P}(X)$  and subsets of  $\mathcal{P}(X)$ :

$$\begin{aligned} \mathbb{G} \subseteq \mathcal{P}(X) \times \mathcal{P}(X) \mapsto \mathbb{G}^\triangleright &:= \{ F \in \mathcal{P}(X) : \forall (Y \rightarrow Z) \in \mathbb{G} (Y \rightarrow Z) \odot F \}, \\ \mathcal{F} \subseteq \mathcal{P}(X) \mapsto \mathcal{F}^\triangleleft &:= \{ (Y \rightarrow Z) \in \mathcal{P}(X) \times \mathcal{P}(X) : \forall F \in \mathcal{F} (Y \rightarrow Z) \odot F \}. \end{aligned}$$

Well-known classic results in lattice theory on Galois connections from [2, §V.7], completed by particular observations from [24, (unnumbered) Theorem], say the following in case of a finite set  $X$ :

- the mapping  $\mathbb{G} \mapsto \mathbb{G}^{\triangleright\triangleleft}$  is a closure operation on subsets of the set  $\mathcal{P}(X) \times \mathcal{P}(X)$  and  $\mathbb{G}^{\triangleright\triangleleft}$  is the smallest functional dependence relation (on subsets of  $X$ ) containing  $\mathbb{G}$ ,
- the mapping  $\mathcal{F} \mapsto \mathcal{F}^{\Phi}$  is a closure operation on subsets of  $\mathcal{P}(X)$  and  $\mathcal{F}^{\Phi}$  is the smallest Moore family (of subsets of  $X$ ) containing  $\mathcal{F}$ ,
- a system  $\mathcal{F} \subseteq \mathcal{P}(X)$  is a Moore family if and only if  $[ \exists \mathbb{G} \subseteq \mathcal{P}(X) \times \mathcal{P}(X) : \mathcal{F} = \mathbb{G}^\triangleright ]$ .

We say that  $\mathbb{G} \subseteq \mathcal{P}(X) \times \mathcal{P}(X)$  is an *implicational generator* of a Moore family  $\mathcal{F}$  if  $\mathcal{F} = \mathbb{G}^\triangleright$ . Elementary facts on Galois connections allow one to observe that an equivalent condition is  $\mathcal{F}^\triangleleft = \mathbb{G}^{\triangleright\triangleleft}$ . Hence,  $\mathbb{G}$  is an implicational generator of a Moore family  $\mathcal{F}$  if and only if both  $\mathbb{G} \subseteq \mathcal{F}^\triangleleft$  (soundness) and  $|\mathbb{G}^\triangleright| = |\mathcal{F}|$  (completeness).

**Example 2.13.** Consider  $X := \{x, y, z, w\}$  and put

$$\mathbb{G} := \{ \emptyset \rightarrow \{x\}, \{z\} \rightarrow \{y\}, \{x, w\} \rightarrow \{x, y, w\} \}.$$

Then  $\mathbb{G}^\triangleright$  equals to the family  $\mathcal{F}$  from Example 2.8. Hence,  $\mathbb{G}$  is one of several possible implicational generators of  $\mathcal{F}$ , a minimal one in sense of inclusion (= a non-redundant one).

To introduce a prominent implicational generator one needs specific concepts from the theory of formal concept analysis; see [15, Chapter 3] or [40, § 1.3.2].<sup>5</sup>

**Definition 2.14** (*Pseudo-Closed Sets, Canonical Basis*).

Given a finite set  $X$ , a subset  $Q \subseteq X$  is *pseudo-closed* relative to a Moore family  $\mathcal{F} \subseteq \mathcal{P}(X)$  if

- $Q \notin \mathcal{F}$  (=  $Q$  is not closed),
- $\forall R \subset Q$  pseudo-closed  $\text{cl}_{\mathcal{F}}(R) \subseteq Q$  (then  $\text{cl}_{\mathcal{F}}(R) \subset Q$  because  $Q \notin \mathcal{F}$ ).

The class of pseudo-closed subsets relative to  $\mathcal{F}$  will be denoted by  $\wp(\mathcal{F})$ . The *canonical implicational basis* for a Moore family  $\mathcal{F}$  is as follows:

$$\mathbb{G}_{\mathcal{F}} := \{ (Q \rightarrow \text{cl}_{\mathcal{F}}(Q) \setminus Q) : Q \in \wp(\mathcal{F}) \} \subseteq \mathcal{P}(X) \times \mathcal{P}(X).$$

The recursive definition of pseudo-closedness of a set  $Q$  is well-formed as it depends only on the previously defined pseudo-closedness of strict subsets of  $Q$ . The canonical implicational basis, alternatively named Guigues–Duquenne basis, was introduced already in the 1980s [18]. We have also implemented an algorithm to compute the canonical basis  $\mathbb{G}_{\mathcal{F}}$  for a given Moore family  $\mathcal{F}$ ; its idea is described in Appendix A.

Every inclusion-minimal non-closed subset, that is,  $Q \in (\mathcal{P}(X) \setminus \mathcal{F})_{\min}$ , is always pseudo-closed since all proper subsets of  $Q$  are closed. The converse implication does not hold as the following example shows.

**Example 2.15.** Consider again the Moore family  $\mathcal{F}$  from Example 2.8 and its closure operation  $\text{cl}_{\mathcal{F}}$ . Then one has  $\wp(\mathcal{F}) = \{ \emptyset, \{x, z\}, \{x, w\} \}$ . Note in this context that none of the singleton subsets of  $X$  is pseudo-closed because

<sup>4</sup> A one-to-one assignment between Moore families and functional dependence relations has already been established through their shared closure operations. We now re-interpret this as a kind of duality relationship and avoid the closure operations.

<sup>5</sup> In general, there is no unique smallest implicational generator in sense of inclusion, that is, a non-redundant generator.



of  $\text{cl}_{\mathcal{F}}(\emptyset) = \{x\}$ . The empty set  $\emptyset$  is the only inclusion-minimal non-closed set while  $\{x, z\}$  and  $\{x, w\}$  are additional pseudo-closed sets. Specifically, one has

$$\mathbb{G}_{\mathcal{F}} = \{ \emptyset \rightarrow \{x\}, \{x, z\} \rightarrow \{y\}, \{x, w\} \rightarrow \{y\} \}.$$

Nevertheless, further simplification is possible. Because of  $\emptyset \rightarrow \{x\}$ , the implication  $\{x, w\} \rightarrow \{y\}$  is equivalent to  $\{w\} \rightarrow \{y\}$  (use axioms of functional dependence relations to observe that). To summarize: the simplest implicational generator for the family  $\mathcal{F}$  seems to be the following list:

$$\{ \emptyset \rightarrow \{x\}, \{z\} \rightarrow \{y\}, \{w\} \rightarrow \{y\} \}.$$

The widely-known result in formal concept analysis says that  $(\mathbb{G}_{\mathcal{F}})^{\circ} = \mathcal{F}$  for any Moore family  $\mathcal{F} \subseteq \mathcal{P}(X)$ ; see [15, Theorem 7 in § 3.2] for a relatively short proof. In fact,  $\mathbb{G}_{\mathcal{F}}$  is a non-redundant implicational generator of  $\mathcal{F}$  in sense that the removal of any of its elements leads to a non-generator.

Let us introduce further specific terminology concerning formal implications. Given a Moore family  $\mathcal{F} \subseteq \mathcal{P}(X)$ , a *valid implication for  $\mathcal{F}$*  is  $(Y \rightarrow Z) \in \mathcal{F}^{\circ}$ . An implication  $Y \rightarrow Z$  is *perfect* (for  $\mathcal{F}$ ) if  $Y \notin \mathcal{F}$  but every strict subset of  $Y$  is closed, that is,  $Y \in (\mathcal{P}(X) \setminus \mathcal{F})_{\min}$ , and  $Z = \text{cl}_{\mathcal{F}}(Y) \setminus Y$ . The implication is *canonical* if it belongs to the canonical implicational basis  $\mathbb{G}_{\mathcal{F}}$ .

It is immediate that every perfect implication is canonical but the converse does not hold as shown in Example 2.15, where  $\{x, z\} \rightarrow \{y\}$  is a canonical implication which is not perfect.

**Definition 2.16** (*Implicatively Perfect Moore Family*).

We say that a Moore family  $\mathcal{F} \subseteq \mathcal{P}(X)$  is *implicatively perfect* if the set of perfect implications for  $\mathcal{F}$  is an implicational generator of  $\mathcal{F}$ .

It makes no problem to deduce from non-redundancy of the canonical basis that an equivalent condition is that every pseudo-closed set is an inclusion-minimal non-closed set, that is, an element of set collection  $(\mathcal{P}(X) \setminus \mathcal{F})_{\min}$ . Let us illustrate the above defined notions by an example.

**Example 2.17** (*Probabilistic Models Over 3 Variables Revisited*).

Consider again the situation from Example 2.10:  $N := \{a, b, c\}$ ,  $X := \mathcal{K}(N)$ , and  $\mathcal{F} := \text{pr}(N)$ . There are six pseudo-closed sets relative to  $\mathcal{F}$  belonging to one permutational type: the set  $\{ab|\emptyset, ac|b\}$  and five of its permutational copies. Each of these sets is minimal in  $\mathcal{P}(X) \setminus \mathcal{F}$ ; thus,  $\mathcal{F}$  is implicatively perfect. The canonical basis  $\mathbb{G}_{\mathcal{F}}$  consists of permutational copies of the implication

$$\{ab|\emptyset, ac|b\} \rightarrow \{ac|\emptyset, ab|c\}.$$

The characterization of this particular Moore family  $\mathcal{F}$  in terms of its canonical implicational basis is equivalent to a later definition of a semi-graphoid over  $\{a, b, c\}$  in Section 3.1.

**Remark.** The concept from Definition 2.16 can be viewed as a counter-part of the one of a coatomistic lattice in the context of implicational description of the lattice. Realize that the Galois connection  $\mathcal{F} \mapsto \mathcal{F}^{\circ}$ , for Moore families  $\mathcal{F}$  of subsets of a finite set  $X$ , can be viewed as an anti-isomorphism between such Moore families and abstract functional dependence relations  $\mathbb{G}$ , whose inverse is the Galois connection  $\mathbb{G} \mapsto \mathbb{G}^{\circ}$ . We can interpret this map as a kind of self-inverse duality transformation between these mathematical structures, which both can be viewed as alternative (ways of) descriptions of finite lattices. If one prefers the perspective of a Moore family then the simplest lattice is the coatomistic one but if one prefers the perspective of implicational description then the simplest lattice is the one described by perfect implications.

2.3. *Frame-based closure of a CI model and implications between CI models*

Let us come back to the situation in Section 2.1 and consider an abstract CI frame  $\mathfrak{f}$  closed under intersection. In this case, for every finite variable set  $N$ , the family of CI models  $(\mathfrak{f}(N), \subseteq)$  is a Moore family of subsets of  $X := \mathcal{K}(N)$ . Therefore, the closure operation  $\text{cl}_{\mathfrak{f}(N)}$  on subsets of  $\mathcal{K}(N)$ , respectively the collection of *valid implications for  $\mathfrak{f}(N)$* , characterizes this family. Observe that, for  $\mathcal{M}, \mathcal{R} \subseteq \mathcal{K}(N)$ , the validity of the implication  $\mathcal{M} \rightarrow \mathcal{R}$  for  $\mathfrak{f}(N)$  reduces to the condition

$$\forall \mathcal{E} \in \mathfrak{f}(N) \quad \mathcal{M} \subseteq \mathcal{E} \text{ implies } \mathcal{R} \subseteq \mathcal{E}.$$

In the sequel, we intend to switch freely between the description of a CI frame  $\mathfrak{f}$  in terms of the families  $\mathfrak{f}(N)$  and in terms of the corresponding closure operations.

The fact that the elements of the canonical basis in Example 2.17 fall down into permutational types is not a surprise because we have the following.

**Observation 2.18.** *If a CI frame  $\mathfrak{f}$  is closed under copying and intersection then, for each set  $N$ , the valid implications  $\mathcal{M} \rightarrow \mathcal{R}$  for the Moore family  $\mathcal{F} := \mathfrak{f}(N)$  and the elements of the canonical implicational basis for  $\mathfrak{f}(N)$  fall into permutational equivalence classes.*

**Proof.** The arguments are analogous to those in [Observation 2.9](#). Any permutation  $\pi : N \rightarrow N$  defines a simultaneous order automorphism of the posets  $(\mathfrak{f}(N), \subseteq)$  and  $(\mathcal{P}(\mathcal{K}(N)), \subseteq)$ . Because such an order automorphism of  $(\mathcal{F}, \subseteq)$  commutes with the closure operation  $\text{cl}_{\mathcal{F}}$ , it preserves the validity of implications.  $\square$

We typically wish to have the frame-based closure stable relative to an enlargement of the basic set.

**Observation 2.19.** *Let  $\mathfrak{f}$  be a CI frame closed under marginalization, intersection, and ascetic extension. Given  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $N \subseteq M$  one has  $\text{cl}_{\mathfrak{f}(N)}(\mathcal{M}) = \text{cl}_{\mathfrak{f}(M)}(\mathcal{M})$ .*

**Proof.** The fact that  $\mathfrak{f}$  is closed under marginalization implies  $\text{cl}_{\mathfrak{f}(N)}(\mathcal{M}) \subseteq \text{cl}_{\mathfrak{f}(M)}(\mathcal{M})$ . Since  $\mathfrak{f}(N) \subseteq \mathfrak{f}(M)$  (ascetic extension), one has  $\text{cl}_{\mathfrak{f}(M)}(\mathcal{M}) \subseteq \text{cl}_{\mathfrak{f}(N)}(\mathcal{M})$ , which yields the equality.  $\square$

Therefore, the assumption on a CI frame  $\mathfrak{f}$  being closed under ascetic extension makes the following definition of the (frame-based)  $\mathfrak{f}$ -closure of a CI model  $\mathcal{M}$  consistent, because the closure does not depend on the “context”, by which we mean a variable set  $M$  such that  $\mathcal{M} \subseteq \mathcal{K}(M)$ .

**Definition 2.20** (*Frame-Based Closure of a CI Model*).

Consider a CI frame  $\mathfrak{f}$  closed under marginalization, intersection, and ascetic extension. Given a CI model  $\mathcal{M} \subseteq \mathcal{K}(N)$  over a variable set  $N$ , its closure of relative to  $\mathfrak{f}(N)$  will be denoted by  $\mathfrak{f}(\mathcal{M}) := \text{cl}_{\mathfrak{f}(N)}(\mathcal{M})$ .

We will call it the  $\mathfrak{f}$ -closure of the model  $\mathcal{M}$ .

Indeed, by [Observation 2.19](#), if  $\mathcal{M} \subseteq \mathcal{K}(N)$  then both  $\mathcal{M}$  and  $\mathfrak{f}(\mathcal{M})$  are models over  $N^{\mathcal{M}} := \bigcup_{ij|K \in \mathcal{M}} ijK$ . Also, if the assumptions of [Definition 2.20](#) are fulfilled and  $\mathcal{M} \subseteq \mathcal{K}(N)$ ,  $\mathcal{R} \subseteq \mathcal{K}(M)$  are given then the condition  $\mathcal{M} \rightarrow \mathcal{R}$  for  $\mathfrak{f}(NM)$  enforces  $N^{\mathcal{R}} \subseteq N^{\mathcal{M}}$  and its verification reduces to showing that  $\mathcal{R} \subseteq \mathfrak{f}(\mathcal{M})$ .

### 3. Examples of abstract CI frames

This is an overview of CI frames; some of these were topics of our computational experiments.

#### 3.1. Semi-graphoids

The concept of a semi-graphoid was introduced by Pearl [32] with the aim to formalize fundamental properties of probabilistic CI pinpointed earlier by Dawid [10]. Here we give an equivalent definition from [25, §5] adapted to our situation. A *semi-graphoid* over a variable set  $N$  is a subset  $\mathcal{M}$  of  $\mathcal{K}(N)$  satisfying

$$ij|K, i\ell|jK \in \mathcal{M} \Leftrightarrow i\ell|K, ij|\ell K \in \mathcal{M}.$$

Then we put

$$\text{sg}(N) := \{ \mathcal{M} \subseteq \mathcal{K}(N) : \mathcal{M} \text{ is a semi-graphoid over } N \},$$

which defines the *semi-graphoidal* CI frame  $\text{sg}$ .

**Lemma 3.1.**  *$\text{sg}$  is closed under copying, marginalization, intersection, duality, and ascetic extension.*

**Proof.** This basically follows from the above definition of a semi-graphoid.  $\square$

Well-known facts are that  $\text{pr} \subseteq \text{sg}$  and  $\text{sg}(N) = \text{pr}(N)$  if  $|N| \leq 3$ . The simplest example showing that  $\text{sg}(N) \neq \text{pr}(N)$  in case  $|N| \geq 4$  is as follows.

**Example 3.2.** Given  $N := \{a, b, c, d\}$  and  $\mathcal{M} := \{ab|c, ac|d, ad|b\} \in \text{sg}(N)$ , one has  $\mathcal{M} \notin \text{pr}(N)$  by property (E:1) in [37, §V.A] saying that the equivalence

$$[ a \perp\!\!\!\perp b \mid c \ \& \ a \perp\!\!\!\perp c \mid d \ \& \ a \perp\!\!\!\perp d \mid b ] \Leftrightarrow [ a \perp\!\!\!\perp b \mid d \ \& \ a \perp\!\!\!\perp c \mid b \ \& \ a \perp\!\!\!\perp d \mid c ]$$

holds in the context of discrete probabilistic CI structures.

#### 3.2. Structural semi-graphoids

The concept of a structural semi-graphoid was introduced in [34] as a semi-graphoid induced in a particular way by certain subsets of  $\mathbb{Z}^{\mathcal{P}(N)}$ . The original complicated definition was shown in [35, Corollary 5.3] to be equivalent to a much simpler condition that underlies our definition below.

**Definition 3.3** (Supermodular Function).

Given a set function  $m : \mathcal{P}(N) \rightarrow \mathbb{R}$  and arbitrary subsets  $I, J, K \subseteq N$ , we are going to use the following special notation for a particular value:

$$\Delta m \langle I, J|K \rangle := m(IJK) + m(K) - m(IK) - m(JK).$$

The supermodular difference expression for  $ij|K \in \mathcal{K}(N)$  is a special case:

$$\Delta m(ij|K) := \Delta m \langle i, j|K \rangle = m(ijK) + m(K) - m(iK) - m(jK).$$

The function  $m$  is called *supermodular* if  $\Delta m(ij|K) \geq 0$  for any  $ij|K \in \mathcal{K}(N)$ , which requirement is equivalent to  $\Delta m \langle I, J|K \rangle \geq 0$  for pairwise disjoint subsets  $I, J, K \subseteq N$ , or to the well-known standard supermodularity condition

$$m(A \cup B) + m(A \cap B) \geq m(A) + m(B) \quad \text{for any } A, B \subseteq N.$$

The model  $\mathcal{M}^m := \{ij|K \in \mathcal{K}(N) : \Delta m(ij|K) = 0\}$ , where  $m : \mathcal{P}(N) \rightarrow \mathbb{R}$  is supermodular, is the structural semi-graphoid induced by  $m$ . The class of *structural semi-graphoids* over  $N$  is

$$\text{st}(N) := \{\mathcal{M}^m : m \text{ is a supermodular function on } \mathcal{P}(N)\}.$$

This defines the structural CI frame  $\text{st}$ .

**Lemma 3.4.**  $\text{st}$  is closed under copying, marginalization, intersection, duality, and ascetic extension.

**Proof.** Given a bijection  $\varphi : N \rightarrow M$  and  $m : \mathcal{P}(N) \rightarrow \mathbb{R}$ , one has  $\Delta m(ij|K) = \Delta(m \circ \varphi_{-1})\varphi(ij|K)$  for any  $ij|K \in \mathcal{K}(N)$ , where  $\varphi$  is naturally extended to the power set of  $N$  and  $\mathcal{K}(N)$ . Hence, provided that  $m$  is a supermodular function on  $\mathcal{P}(N)$ , the composition  $m \circ \varphi_{-1}$  is a supermodular function on  $\mathcal{P}(M)$  and the  $\varphi$ -copy of  $\mathcal{M}^m$  is  $\mathcal{M}^{m \circ \varphi_{-1}}$ .

If  $M \subseteq N$  then the restriction of a supermodular function on  $\mathcal{P}(N)$  to  $\mathcal{P}(M)$  is supermodular, which implies the claim about the marginalization.

Given two set functions  $m^1 : \mathcal{P}(N) \rightarrow \mathbb{R}$  and  $m^2 : \mathcal{P}(N) \rightarrow \mathbb{R}$ , one has  $\Delta(m^1 + m^2)(ij|K) = \Delta m^1(ij|K) + \Delta m^2(ij|K)$  for any  $ij|K \in \mathcal{K}(N)$ , which implies that if  $m^1$  and  $m^2$  are supermodular then  $m := m^1 + m^2$  is supermodular. The non-negativity of values  $\Delta m^1(ij|K)$  and  $\Delta m^2(ij|K)$  allows one to deduce that  $\mathcal{M}^m = \mathcal{M}^{m^1} \cap \mathcal{M}^{m^2}$ .

Consider a self-inverse mapping  $\iota : \mathcal{P}(N) \rightarrow \mathcal{P}(N)$  defined by  $\iota(S) := N \setminus S$  for any  $S \subseteq N$ . The formula  $\Delta m(ij|K) = \Delta(m \circ \iota)(ij|K)^*$  valid for any  $m : \mathcal{P}(N) \rightarrow \mathbb{R}$  and  $ij|K \in \mathcal{K}(N)$ , implies that  $m$  is supermodular if and only if  $m \circ \iota$  is supermodular and that  $(\mathcal{M}^m)^* = \mathcal{M}^{m \circ \iota}$ .

To show that  $\text{st}$  is closed under extension, realize that, if  $N \subseteq M$ , then any supermodular function  $r : \mathcal{P}(N) \rightarrow \mathbb{R}$  can be extended to a supermodular function  $m$  on  $\mathcal{P}(M)$  by the formula  $m(T) := r(T \cap N)$  for any  $T \subseteq M$ . Since  $\text{st}$  is closed under intersection, to show it is closed under ascetic extension it remains to observe that  $\mathcal{K}(N) \in \text{st}(M)$ . This follows from  $\text{pr} \subseteq \text{st}$ ; a specific construction is left to the reader.  $\square$

Basic facts are that  $\text{pr} \subseteq \text{st} \subseteq \text{sg}$  and  $\text{sg}(N) = \text{st}(N) = \text{pr}(N)$  if  $|N| \leq 3$ ; for  $\text{pr} \subseteq \text{st}$  see [35, §2.3.4]. The strict inclusion  $\text{st}(N) \subset \text{sg}(N)$  in case  $|N| = 4$  is evidenced by the model  $\mathcal{M} := \{ab|c, ac|d, ad|b\}$  from Example 3.2, because the property (E:1) mentioned there holds for structural semi-graphoids.

The strict inclusion  $\text{pr}(N) \subset \text{st}(N)$  for  $|N| = 4$  is evidenced by  $\mathcal{M}^* := \{ab|\emptyset, ab|c, ab|d, cd|ab\}$ , which is the dual model in Example 2.6, induced by the next supermodular function  $m$  on  $N = \{a, b, c, d\}$ :

$$m(S) := \begin{cases} 12 & \text{if } S = abcd, \\ 7 & \text{if } S = acd \text{ or } S = bcd, \\ 6 & \text{if } S = abc \text{ or } S = abd, \\ 3 & \text{if } S \in \{ac, ad, bc, bd, cd\}, \\ 0 & \text{otherwise,} \end{cases} \quad \text{for } S \subseteq N.$$

*Remarks.* This is to remind alternative views on structural semi-graphoids.

1. A set function  $h : \mathcal{P}(N) \rightarrow \mathbb{R}$  is *submodular* if it fulfils the converse inequalities to the supermodular ones:

$$h(A \cup B) + h(A \cap B) \leq h(A) + h(B) \quad \text{for any } A, B \subseteq N,$$

which is another way of saying that  $-h$  is supermodular. The rank function of a *polymatroid* over  $N$  is a submodular function  $h$  on  $\mathcal{P}(N)$  which satisfies  $h(\emptyset) = 0$  and is non-decreasing:  $h(A) \leq h(B)$  if  $A \subseteq B \subseteq N$ .

One can equivalently introduce structural semi-graphoids as CI models induced by submodular functions or by polymatroids in an analogous way. Clearly, it is just a matter of taste and convention whether the supermodular view or the submodular view is chosen.

2. The concept of a structural semi-graphoid can be viewed as a reduced version of the concept of a *semi-matroid* from [30, §2]. The difference is that semi-matroids also involve *functional dependencies* within  $N$  (see the remark concluding Section 2.1). But otherwise semi-matroids are defined analogously: they are induced by polymatroids through vanishing of the respective submodular difference expressions.

The polymatroidal attitude allows one to emphasize the links to *matroid theory* [31] because polymatroids generalize matroids. Specifically, a polymatroidal rank function  $h$  on  $\mathcal{P}(N)$  corresponds to a *matroid* over  $N$  if and only if it takes values in  $\mathbb{Z}$  and additionally satisfies  $h(I) \leq |I|$  for any  $I \subseteq N$ .

### 3.3. Graphoids, compositional semi-graphoids, and compositional graphoids

Researchers interested in graphical statistical models introduced various other abstract CI frames [32,33]; here we follow the line from [23]. A *graphoid* over a variable set  $N$  is a semi-graphoid  $\mathcal{M} \subseteq \mathcal{K}(N)$  satisfying the implication

$$ij|\ell K, i\ell|jK \in \mathcal{M} \Rightarrow ij|K, i\ell|K \in \mathcal{M}.$$

A semi-graphoid is called *compositional* if it satisfies the reversed implication, and a *compositional graphoid* is a semi-graphoid satisfying both implications. It is evident that  $\mathcal{M}$  is a graphoid if and only if its dual model  $\mathcal{M}^*$  is a compositional semi-graphoid.

The above definitions, thus, delimit the frame  $\text{gr}$  of graphoids, the frame  $\text{gr}^*$  of compositional semi-graphoids, and the frame  $\text{cgr}$  of compositional graphoids. By definition,  $\text{cgr} \subseteq \text{gr}, \text{gr}^* \subseteq \text{sg}$ . Our notation reflects the fact that  $\text{gr}^*$  can be interpreted as the dual frame to  $\text{gr}$  because  $\mathcal{M} \in \text{gr}^*(N) \Leftrightarrow \mathcal{M}^* \in \text{gr}(N)$  for any variable set  $N$  and  $\mathcal{M} \subseteq \mathcal{K}(N)$ . It is known that  $\text{gr} \setminus \text{pr} \neq \emptyset$  and  $\text{pr} \setminus \text{gr} \neq \emptyset$ .

### 3.4. Graphical CI frames

A quite special CI frame is given by undirected graphical models, which were called *separation graphoids* in [23]. Given an undirected graph  $G$  over a variable set  $N$  (= a graph having  $N$  as the node set), we say that  $ij|K \in \mathcal{K}(N)$  is represented in  $G$  (after the separation criterion) and write  $i \perp\!\!\!\perp j \mid K [G]$  if every path in  $G$  between  $i$  and  $j$  contains a node in  $K$ . Then  $G$  induces a CI model

$$\mathcal{M}_G := \{ij|K : i \perp\!\!\!\perp j \mid K [G]\} \subseteq \mathcal{K}(N).$$

The family of *undirected graphical models* is then

$$\text{ug}(N) := \{\mathcal{M}_G : G \text{ is an undirected graph over } N\}.$$

This defines a CI frame  $\text{ug}$ , which is closed under copying and marginalization, but not under intersection. Well-known facts are that  $\text{ug} \subseteq \text{pr} \subseteq \text{sg}$  [17]. The  $\text{ug}$ -frame has a finite “axiomatization” [32].

A popular class of CI models studied in probabilistic reasoning are *directed acyclic graphical models*. These are induced by directed graphs without directed cycles through the respective  $d$ -separation criterion [32]. The corresponding CI frame is closed under copying but not under marginalization [35, Example 3.1].

### 3.5. Other examples of abstract CI frames

Another CI frame, which is not discussed in this paper, consists of *regular Gaussian CI models* [3], that is, CI structures induced by (continuous) Gaussian probability distributions. This frame is closed under copying, marginalization, and duality but not under intersection.

The frame of *gaussoids* was introduced in [23] and examined in [5]. This frame approximates the frame of (regular) Gaussian CI models. A recent paper [4] deals with the structural self-adhesivity concept for gaussoids and applies these techniques to Gaussian random variables.

Further options were pinpointed by one of the reviewers. One can consider the *matroidal frame*, a subframe of  $\text{st}$  consisting of CI models induced by rank functions of matroids. The other option is the *linear frame*, consisting of CI models induced by linear polymatroids. Yet another option is the *almost entropic frame*, discussed in more details in Section 8.

## 4. Adhesion and self-adhesion

The concept of an adhesion/self-adhesion was introduced by Matúš [29] in the context of polymatroids. As pinpointed by Csirmaz [9], the adhesion concept is a special case of the concept of an amalgam for polymatroids. Note that amalgams were introduced and studied much earlier in matroid theory [31].

Matúš’s concept of self-adhesion can be viewed as an abstraction of the procedure used to derive non-Shannon information inequalities [43], which is commonly named the “*copy lemma*” method [13] in information-theoretical community. More details on the connection to this field are discussed in Section 8.

It has been observed recently in [3] that these concepts can also be defined within a simpler discrete environment of CI structures. We follow this link here. An auxiliary concept of consonant models is needed.

**Definition 4.1** (*Consonant Models, Amalgam, Adhesion*).

- Let  $N, M$  be two variable sets, possibly intersecting. Models  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $\mathcal{R} \subseteq \mathcal{K}(M)$  will be called *consonant* if  $\mathcal{M} \downarrow^{N \cap M} = \mathcal{R} \downarrow^{N \cap M}$ .
- Given consonant models  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $\mathcal{R} \subseteq \mathcal{K}(M)$ , their *amalgam* is any “extended” model  $\mathcal{E} \subseteq \mathcal{K}(NM)$  over  $NM$  such that  $\mathcal{E} \downarrow^N = \mathcal{M}$  and  $\mathcal{E} \downarrow^M = \mathcal{R}$ .

- An *adhesion* of  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $\mathcal{R} \subseteq \mathcal{K}(M)$  is their amalgam  $\mathcal{E} \subseteq \mathcal{K}(NM)$  that, moreover, satisfies  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid (N \cap M) [\mathcal{E}]$  (see Definition 2.3).

Note that any pair of models  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $\mathcal{R} \subseteq \mathcal{K}(M)$  is consonant if  $|N \cap M| \leq 1$ . The consonancy is a necessary condition for the existence of an amalgam. Observe that any pair of consonant models has an adhesion: take the union of  $\mathcal{M}$ ,  $\mathcal{R}$ , and the subset of  $\mathcal{K}(NM)$  delimiting the global CI statement  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid (N \cap M)$ .

Unlike the algebraic operations from Definition 2.4, amalgams and adhesions are not uniquely determined by their input models  $\mathcal{M}$  and  $\mathcal{R}$ . On the other hand, (set-theoretical) intersection of two amalgams/adhesions of  $\mathcal{M}$  and  $\mathcal{R}$  is their amalgam/adhesion; thus, their least amalgam/adhesion exists.

**Example 4.2.** Take  $N := \{a, b, c, d\}$ ,  $M := \{c, d, e\}$ , and put  $\mathcal{M} := \{ab|\emptyset, cd|\emptyset\}$ ,  $\mathcal{R} := \{cd|\emptyset, cd|e\}$ , and  $\mathcal{R}' := \{cd|e\}$ . Then  $\mathcal{M}$  and  $\mathcal{R}$  are consonant models, while  $\mathcal{M}$  and  $\mathcal{R}'$  are not. The least amalgam of  $\mathcal{M}$  and  $\mathcal{R}$  is  $\{ab|\emptyset, cd|\emptyset, cd|e\}$  and the least adhesion of  $\mathcal{M}$  and  $\mathcal{R}$  is

$$\{ab|\emptyset, cd|\emptyset, cd|e, ae|cd, ae|bcd, be|cd, be|acd\},$$

where the last 4 statements represent the global CI statement  $ab \perp\!\!\!\perp e \mid cd$ .

#### 4.1. Self-adhesion

The self-adhesion of a CI model  $\mathcal{M}$  is nothing but the adhesion of  $\mathcal{M}$  with its own  $\varphi$ -copy, where  $\varphi$  is an appropriate bijection. The derived concept of self-adhesivity is relative to a CI frame.

**Definition 4.3 (Self-Adhesion).**

- Let  $\mathfrak{f}$  be a CI frame closed under copying and marginalization. Given  $\mathcal{M} \in \mathfrak{f}(N)$  and  $L \subseteq N$  we say that  $\mathcal{M}$  is *self-adhesive at  $L$*  relative to  $\mathfrak{f}$  if, for every bijection  $\varphi : N \rightarrow M$  such that  $L = N \cap M$  and  $\varphi|_L$  is the identity, an adhesion of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$  exists which is in the family  $\mathfrak{f}(NM)$ :

$$\forall \varphi : N \rightarrow M \text{ bijection with } L = N \cap M \text{ and } \varphi|_L = \text{id}_L \quad \exists \mathcal{E} \in \mathfrak{f}(NM) \\ \text{such that } \mathcal{E} \downarrow^N = \mathcal{M}, \mathcal{E} \downarrow^M = \varphi(\mathcal{M}) \text{ and } (N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid L [\mathcal{E}].$$

We say that  $\mathcal{M} \in \mathfrak{f}(N)$  is *self-adhesive* relative to  $\mathfrak{f}$  if it is self-adhesive at every  $L \subseteq N$ .

- The symbol  $\mathfrak{f}^\circ(N)$  will denote the family of *self-adhesive models*  $\mathcal{M} \in \mathfrak{f}(N)$  relative to  $\mathfrak{f}$ . If  $\mathfrak{f}(N) = \mathfrak{f}^\circ(N)$  then we say that  $\mathfrak{f}(N)$  is self-adhesive. The frame  $\mathfrak{f}$  is *self-adhesive* if this holds for every finite  $N$ .
- Given  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $L \subseteq N$ , if there exists the least model  $\mathcal{R} \in \mathfrak{f}(N)$  such that  $\mathcal{M} \subseteq \mathcal{R}$  and  $\mathcal{R}$  is self-adhesive at  $L$  relative to  $\mathfrak{f}$  then we will call  $\mathcal{R}$  the *self-adhesive closure* of  $\mathcal{M}$  at  $L$  (relative to  $\mathfrak{f}$ ) and denote it by  $\mathfrak{f}^\circ(\mathcal{M}|L)$ .

The assumptions on  $\mathfrak{f}$  (copying and marginalization) are needed to define the concept of self-adhesion. Clearly, a model  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at  $L \subseteq N$  relative to  $\mathfrak{f}$  if and only if  $\mathfrak{f}^\circ(\mathcal{M}|L) = \mathcal{M}$ .

*Remarks.* We have a few additional related remarks.

1. One can equivalently introduce a self-adhesive model  $\mathcal{M}$  at a set  $L$  by a formally weaker condition:

$$\exists \psi : N \rightarrow O \text{ bijection with } L = N \cap O \text{ and } \psi|_L = \text{id}_L \quad \exists \tilde{\mathcal{E}} \in \mathfrak{f}(NO) : \\ \tilde{\mathcal{E}} \downarrow^N = \mathcal{M} \ \& \ \tilde{\mathcal{E}} \downarrow^O = \psi(\mathcal{M}) \ \& \ (N \setminus O) \perp\!\!\!\perp (O \setminus N) \mid L [\tilde{\mathcal{E}}].$$

This follows from the assumption that  $\mathfrak{f}$  is closed under copying. Having fixed  $\psi : N \rightarrow O$ , for any considered bijection  $\varphi : N \rightarrow M$  we define  $\rho : NO \rightarrow NM$  by  $\rho(i) := i$  for  $i \in N$  and  $\rho(i) := \varphi(\psi_{-1}(i))$  for  $i \in O \setminus N$ . Then  $\mathcal{E} := \rho(\tilde{\mathcal{E}})$  is the desired adhesion for  $\varphi$ . Therefore, the choice of the bijection  $\varphi$  is irrelevant and the validity of the self-adhesivity condition is, in fact, an intrinsic property of  $\mathcal{M}$  (relative to  $\mathfrak{f}$ ).

2. This is to explain why the self-adhesivity concept is relative to a CI frame. Having two frames  $\mathfrak{f}_1 \subseteq \mathfrak{f}_2$ , imagine one finds a model  $\mathcal{M} \in \mathfrak{f}_1(N)$  and a bijection  $\varphi : N \rightarrow M$  with  $L = N \cap M$  such that a self-adhesion  $\mathcal{E} \in \mathfrak{f}_2(NM)$  of  $\mathcal{M}$  for  $\varphi$  exists while there is no such a self-adhesion within  $\mathfrak{f}_1(NM)$ . Then  $\mathcal{M}$  is self-adhesive at  $L$  relative to  $\mathfrak{f}_2$  but not relative to  $\mathfrak{f}_1$  (see Example 4.4).
3. It follows from the definition that, given  $\mathcal{M} \in \mathfrak{f}(N)$ , the question whether  $\mathcal{M} \in \mathfrak{f}^\circ(N)$  can be answered on basis of the characterization/description of families  $\mathfrak{f}(M)$ , where  $|N| \leq |M| \leq 2 \cdot |N|$ . Nevertheless, further reduction is possible. For example, it is immediate that any  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at  $L := N$ . Additional reductions, based on further assumptions on the frame, are discussed below.

**Example 4.4 (Self-Adhesivity is Relative to a Frame).**

Consider frames  $\mathfrak{f}_1 := \text{ug}$ ,  $\mathfrak{f}_2 := \text{sg}$  and the variable set  $N := \{a, b, c\}$ . The model  $\mathcal{M} := \{bc|a\} \subseteq \mathcal{K}(N)$  belongs to  $\text{ug}(N)$ : it is induced by the graph  $G$  over  $N$  with two edges, those connecting  $a$  with the other two nodes. Then  $\mathcal{M} \in \text{sg}^\circ(N)$  but  $\mathcal{M} \notin \text{ug}^\circ(N)$ . To this end consider the set  $L := bc$  and a bijection  $\varphi : N \rightarrow M := \{\bar{a}, b, c\}$  with  $\varphi(a) := \bar{a}$ ,  $\varphi(b) := b$ , and

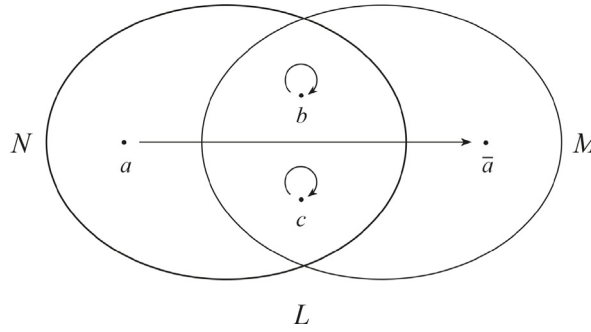


Fig. 2. The bijection  $\varphi : N \rightarrow M$  considered in Examples 4.4 and 4.6.

$\varphi(c) := c$  (see Fig. 2). Then  $\mathcal{M}$  is not self-adhesive at  $L$  relative to  $\text{ug}$ . Indeed, a potential adhesion  $\mathcal{E}$  of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$  in  $\text{ug}(NM)$  has to satisfy

$$\{bc|a, bc|\bar{a}, a\bar{a}|bc\} \subseteq \mathcal{E} \text{ and } \{ab|c, ac|b, \bar{a}b|c, \bar{a}c|b\} \cap \mathcal{E} = \emptyset,$$

and there is no undirected graph  $H$  over  $NM$  with  $\mathcal{M}_H = \mathcal{E}$ .

If a CI frame  $\mathfrak{f}$  is closed under *intersection* then the self-adhesive closure of any model at any set is easy to compute by applying the  $\mathfrak{f}$ -closure. In particular, the question of self-adhesivity testing can be simplified.

**Lemma 4.5.** *Let  $\mathfrak{f}$  be a CI frame closed under copying, marginalization, intersection, and ascetic extension. Given  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $L \subseteq N$ , the self-adhesive closure  $\mathcal{R} = \mathfrak{f}^\circ(\mathcal{M}|L)$  exists and can be computed as follows. Having fixed a bijection  $\varphi : N \rightarrow M$  with  $L = N \cap M$  and  $\varphi|_L = \text{id}_L$ , apply the next steps:*

1. take the least self-adhesion of  $\mathcal{M}$ :  $\mathcal{E} := \mathcal{M} \cup \varphi(\mathcal{M}) \cup [N \setminus M, M \setminus N|L]$ ,
2. apply the  $\mathfrak{f}$ -closure to it:  $\tilde{\mathcal{E}} := \mathfrak{f}(\mathcal{E})$ , (see Definition 2.20)
3. apply marginalization to it:  $\mathcal{R} := \tilde{\mathcal{E}}^{\downarrow N}$ .

**Proof.** Extend the bijection  $\varphi : N \rightarrow M$  by the restriction of its own inversion  $\varphi_{-1}$  to  $M \setminus N$  to get a permutation  $\psi$  of  $NM$ . Observe that  $\psi(\mathcal{E}) = \mathcal{E}$ . Because  $\mathfrak{f}$  is closed under copying, by arguments in Observation 2.7, the  $\mathfrak{f}$ -closure over  $NM$  commutes with  $\psi$  and one even has  $\psi(\tilde{\mathcal{E}}) = \tilde{\mathcal{E}}$ . It follows from the construction and the fact that  $\mathfrak{f}$  is closed under marginalization that  $\mathcal{M} \subseteq \mathcal{R} \in \mathfrak{f}(N)$ . Using that  $\tilde{\mathcal{E}}$  is fixed by  $\psi$  and the fact that copying commutes with marginalization, one has

$$\tilde{\mathcal{E}}^{\downarrow M} = \tilde{\mathcal{E}}^{\downarrow \psi(N)} = \psi(\tilde{\mathcal{E}})^{\downarrow \psi(N)} = \psi|_N(\tilde{\mathcal{E}}^{\downarrow N}) = \varphi(\tilde{\mathcal{E}}^{\downarrow N}) = \varphi(\mathcal{R}),$$

which implies, by Definition 4.3, that  $\mathcal{R}$  is self-adhesive at  $L$ .

Given  $\mathcal{R}' \in \mathfrak{f}(N)$  which is self-adhesive at  $L$ , and satisfies  $\mathcal{M} \subseteq \mathcal{R}'$ , it is easy to observe that the respective adhesion  $\mathcal{E}' \in \mathfrak{f}(NM)$  satisfies  $\mathcal{E} \subseteq \mathcal{E}'$ , which yields  $\tilde{\mathcal{E}} \subseteq \mathcal{E}'$  and, hence,  $\mathcal{R} \subseteq \mathcal{R}'$ . Therefore, one indeed has  $\mathcal{R} = \mathfrak{f}^\circ(\mathcal{M}|L)$ .  $\square$

Note that the  $\mathfrak{f}$ -closure is applied only once and it is over a variable set of cardinality  $|N| + |L|$ . The next example illustrates the method from Lemma 4.5.

**Example 4.6.** Take  $N := \{a, b, c\}$  and put  $\mathcal{M} := \{ab|c\}$ ; note that  $\mathcal{M} \in \text{sg}(N)$ . We are going to test self-adhesivity of  $\mathcal{M}$  at the set  $L := \{b, c\}$  relative to the semi-graphoidal frame  $\text{sg}$ . As explained (in the remark n. 1 above) it suffices to fix a variable set  $M := \{\bar{a}, b, c\}$  and a bijection  $\varphi : N \rightarrow M$  defined by  $\varphi(a) := \bar{a}$ ,  $\varphi(b) := b$ , and  $\varphi(c) := c$ , (see Fig. 2) and to find an adhesion of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$  which belongs to  $\text{sg}(NM)$ . To this end we consider their least adhesion  $\mathcal{E} = \{ab|c, \bar{a}b|c, a\bar{a}|bc\}$  and take its semi-graphoidal closure

$$\text{sg}(\mathcal{E}) = \{ab|c, ab|\bar{a}c, a\bar{a}|c, a\bar{a}|bc, \bar{a}b|c, \bar{a}b|ac\}.$$

Since  $\text{sg}(\mathcal{E})^{\downarrow N} = \mathcal{M}$ , one has  $\text{sg}^\circ(\mathcal{M}|L) = \mathcal{M}$ , and  $\mathcal{M}$  is self-adhesive at  $L$  relative to  $\text{sg}$ . Observe that one has  $\text{sg}(\mathcal{E})^{\downarrow M} = \{\bar{a}b|c\} = \varphi(\mathcal{M})$ .

This was a positive example of self-adhesivity at a set. On the other hand, Example 7.1 (in Section 7.3), obtained as a result of our computations, shows that it may happen that  $\text{sg}^\circ(\mathcal{M}|L) \neq \mathcal{M}$ .

**Lemma 4.7.** *The discrete probabilistic CI frame  $\text{pr}$  is self-adhesive:  $\text{pr}^\circ = \text{pr}$ .*



**Proof.** Given  $\mathcal{M} \in \text{pr}(N)$ , there exists a discrete random vector  $\xi$  over  $N$  such that  $\mathcal{M}$  is induced by  $\xi$ :  $\mathcal{M}_\xi = \mathcal{M}$ . Given a bijection  $\varphi : N \rightarrow M$  with  $\varphi|_{N \cap M}$  being the identity, define a random vector  $\eta := [\xi_{\varphi^{-1}(j)}]_{j \in M}$  over  $M$  and have  $\mathcal{M}_\eta = \varphi(\mathcal{M})$ . Then  $\xi$  and  $\eta$  completely coincide on  $N \cap M$  in the sense that  $\xi_{N \cap M} = \eta_{N \cap M}$ . Hence, one can construct a random vector  $\zeta$  over  $NM$  with  $\zeta_N = \xi$ ,  $\zeta_M = \eta$ , and  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid (N \cap M) [\zeta]$ . This step corresponds to a well-known construction of the *conditional product* of two consonant multi-dimensional measures; see [35, §2.3.3]. Thus, the model  $\mathcal{E} := \mathcal{M}_\zeta$  is the desired adhesion of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$  belonging to the family  $\text{pr}(NM)$ .  $\square$

#### 4.2. Self-adhesivity at small sets: lifting

In next two subsections we show that, if a CI frame is closed under appropriate extending algebraic operations, then the involved models are automatically self-adhesive at sets of particular cardinality.

The first such operation, called *lifting*, ensures the self-adhesivity at the empty set and singletons. It formalizes the act of adding completely independent variables. Specifically, in case of a CI model  $\mathcal{M} \subseteq \mathcal{K}(N)$  and a single (outside) variable  $o \notin N$ , the respective (one-variable) lift is the following model over  $No$ :

$$\mathcal{M} \cup \{ij|Ko : ij|K \in \mathcal{M}\} \cup \{io|K : i \in N, K \subseteq N \setminus i\}.$$

Such one-element (independent) additions commute each other and allow one to introduce the addition of a set of independent variables. Nonetheless, for technical reasons, we prefer the following condensed definition, which is equivalent to the result of repeated one-variable additions.

**Definition 4.8 (Lifting).**

Let  $N \subseteq O$  be variable sets in inclusion and  $\mathcal{M} \subseteq \mathcal{K}(N)$ . The *lift of  $\mathcal{M}$  to  $O$*  is the following model over  $O$ :

$$\mathcal{M}_{N \nearrow O} := \{ij|K \in \mathcal{K}(O) : ij \cap (O \setminus N) \neq \emptyset \text{ or } ij|(K \cap N) \in \mathcal{M}\}.$$

A CI frame  $\mathfrak{f}$  is closed under *lifting*, if for any  $\mathcal{M} \in \mathfrak{f}(N)$  and any superset  $O \supseteq N$ , one has  $\mathcal{M}_{N \nearrow O} \in \mathfrak{f}(O)$ .

Note that lifting complies with marginalization in the sense that, for every  $\mathcal{M} \subseteq \mathcal{K}(N)$  and  $N \subseteq O$ , one has  $(\mathcal{M}_{N \nearrow O})^{\downarrow N} = \mathcal{M}$ .

**Example 4.9.** Put  $N := \{a, b\}$ ,  $O := \{a, b, c\}$ ,  $\mathcal{M} := \emptyset$ , and  $\mathcal{R} := \{ab|\emptyset\}$  and interpret both  $\mathcal{M}$  and  $\mathcal{R}$  as CI models over  $N$ . Then  $\mathcal{M}_{N \nearrow O} = \{ac|\emptyset, ac|b, bc|\emptyset, bc|a\}$ . Note that the variable  $c$  became independent of the rest. In the other case of  $\mathcal{R}$ , one gets the full model:  $\mathcal{R}_{N \nearrow O} = \mathcal{K}(O)$ .

**Lemma 4.10.** *The CI frames  $\text{pr}$ ,  $\text{st}$ , and  $\text{sg}$  are closed under lifting.*

**Proof.** Given a discrete random vector  $\xi = [\xi_i]_{i \in N}$  over  $N$  and  $N \subseteq O$ , one can extend it to a random vector  $\eta = [\eta_i]_{i \in O}$  over  $O$  with stochastically independent sub-vectors  $\eta_N := \xi$  and  $\eta_i$ ,  $i \in O \setminus N$ . Then  $\mathcal{M}_\eta = (\mathcal{M}_\xi)_{N \nearrow O}$ .

Given a supermodular function  $r : \mathcal{P}(N) \rightarrow \mathbb{R}$  and  $N \subseteq O$ , one can extend it to a supermodular function  $m$  on  $\mathcal{P}(O)$  by the formula  $m(S) := r(S \cap N)$  for any  $S \subseteq O$ . It is easy to observe that  $\mathcal{M}^m = (\mathcal{M}^r)_{N \nearrow O}$ .

Given a semi-graphoid  $\mathcal{M} \in \text{sg}(N)$  over  $N$  and  $N \subseteq O$ , it is straightforward to observe that  $\mathcal{M}_{N \nearrow O}$  is a semi-graphoid over  $O$ .  $\square$

The lifting operation is related to adhesions of CI models at small sets.

**Lemma 4.11.** *Let  $\mathfrak{f}$  be a CI frame which is closed under marginalization, intersection, and lifting. Then, for every  $\mathcal{M} \in \mathfrak{f}(N)$  and  $\mathcal{R} \in \mathfrak{f}(M)$  such that  $|N \cap M| \leq 1$ , an adhesion  $\mathcal{E} \in \mathfrak{f}(NM)$  of  $\mathcal{M}$  and  $\mathcal{R}$  exists.*

**Proof.** Put  $\hat{\mathcal{M}} := \mathcal{M}_{N \nearrow NM}$  and observe that  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid (N \cap M) [\hat{\mathcal{M}}]$  and  $\hat{\mathcal{M}}^{\downarrow N} = \mathcal{M}$ . Moreover,  $|N \cap M| \leq 1$  implies  $\hat{\mathcal{M}}^{\downarrow M} = \mathcal{K}(M)$  and the fact  $\hat{\mathcal{M}} \in \mathfrak{f}(NM)$  follows from the assumption that  $\mathfrak{f}$  is closed under lifting. Putting  $\hat{\mathcal{R}} := \mathcal{R}_{M \nearrow NM} \in \mathfrak{f}(NM)$  gives  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid (N \cap M) [\hat{\mathcal{R}}]$  with  $\hat{\mathcal{R}}^{\downarrow N} = \mathcal{K}(N)$  and  $\hat{\mathcal{R}}^{\downarrow M} = \mathcal{R}$ . Since  $\mathfrak{f}$  is closed under intersection,  $\mathcal{E} := \hat{\mathcal{M}} \cap \hat{\mathcal{R}}$  belongs to  $\mathfrak{f}(NM)$  and, moreover, by the above facts,  $\mathcal{E}$  is an adhesion of  $\mathcal{M}$  and  $\mathcal{R}$ .  $\square$

Of course, the above observation on adhesivity applies to self-adhesivity.

**Corollary 4.12.** *Let  $\mathfrak{f}$  be a CI frame which is closed under copying, marginalization, intersection, and lifting. Then, for any  $N$ , any model  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at every set of cardinality 0 or 1.*

**Proof.** Apply Lemma 4.11 to  $\mathcal{M}$  and  $\mathcal{R} := \varphi(\mathcal{M})$  for the appropriate bijection  $\varphi : N \rightarrow M$ .  $\square$

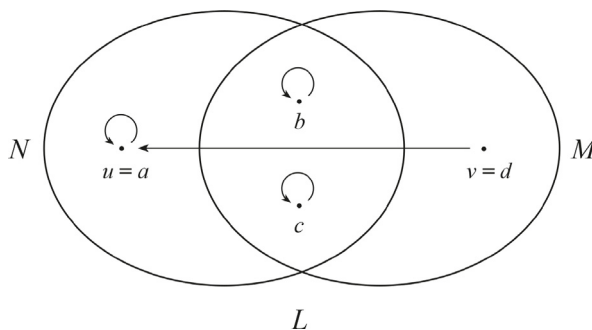


Fig. 3. The auxiliary mapping  $\rho : NM \rightarrow N$  considered in Example 4.14.

### 4.3. Self-adhesivity at co-singletons: tight replication

In this subsection we discuss a particular extending algebraic operation which ensures that, if a CI frame  $\mathfrak{f}$  is closed under this operation, then every  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at *co-singletons*, that is, at the complements of singletons within the variable set  $N$ . This operation is called *tight replication* and formalizes the act of adding one identical copy of a fixed variable in  $N$ . This (technically complex) operation was inspired by an analogous (poly-)matroidal operation of *parallel extension*.

Let us first remind what is our goal and re-formulate the condition of self-adhesivity at a co-singleton (we wish to ensure). Consider  $\mathcal{M} \in \mathfrak{f}(N)$ , fix a variable  $u \in N$  and an outside variable  $v \notin N$ . Consider the co-singleton  $L := N \setminus u$ , denote  $M := Lv$ , and introduce  $\psi : N \rightarrow M$  which is identical on  $L$  and  $\psi(u) = v$ . By Definition 4.3, our goal is to verify the existence of

$$\mathcal{E} \in \mathfrak{f}(NM) \text{ which satisfies } \mathcal{E}^{\downarrow N} = \mathcal{M}, \mathcal{E}^{\downarrow M} = \psi(\mathcal{M}), \text{ and } u \perp\!\!\!\perp v \mid L [\mathcal{E}].$$

One of possible ways is to define  $\mathcal{E}$  as a particular extension of  $\mathcal{M}$  in which  $v$  is the “identical” copy of  $u$ . To this end an auxiliary mapping  $\rho : NM \rightarrow N$  is used, which is an extension of the inverse  $\psi_{-1} : M \rightarrow N$ .

**Definition 4.13 (Tight Replication).**

Let  $N, M$  be two variable sets of cardinality  $n \geq 1$  such that  $|N \cap M| = n - 1$ . Let us denote  $L := N \cap M$ ,  $u := N \setminus L$ ,  $v := M \setminus L$ , and introduce a mapping  $\rho : NM \rightarrow N$  by  $\rho(i) := i$  for  $i \in N$  and  $\rho(v) := u$ . Given  $\mathcal{M} \subseteq \mathcal{K}(N)$ , its *tight replication* (based on  $\rho$ ) is the model  $\mathcal{M}_{v\parallel u} \subseteq \mathcal{K}(NM)$  defined as follows:

- given  $ij|K \in \mathcal{K}(NM)$ ,  $ij|K \in \mathcal{M}_{v\parallel u}$  if and only if  $\rho(i) \perp\!\!\!\perp \rho(j) \mid \rho(K) [\mathcal{M}]$ ,
- under conventions that  $a \perp\!\!\!\perp b \mid C [\mathcal{M}]$  whenever  $ab \cap C \neq \emptyset$
- and  $u \perp\!\!\!\perp u \mid C [\mathcal{M}]$  for  $C \subseteq L$  means  $u \perp\!\!\!\perp (L \setminus C) \mid C [\mathcal{M}]$ .

A CI frame  $\mathfrak{f}$  is closed under *tight replication*, if for any such a pair  $N, M$  of variable sets and  $\mathcal{M} \in \mathfrak{f}(N)$ , one has  $\mathcal{M}_{v\parallel u} \in \mathfrak{f}(NM)$ .

Note that tight replication also complies with marginalization in the sense that  $(\mathcal{M}_{v\parallel u})^{\downarrow N} = \mathcal{M}$ .

**Example 4.14.** Consider the situation from Definition 4.13 when one has  $N := \{a, b, c\}$ ,  $M := \{b, c, d\}$ ,  $L := \{b, c\}$ ,  $u := a$  and  $v := d$ . Define  $\rho(a) := a$ ,  $\rho(b) := b$ ,  $\rho(c) := c$ ,  $\rho(d) := a$  (see Fig. 3 for illustration), put  $\mathcal{M} := \emptyset$  and  $\mathcal{R} := \{ab|c\}$ . Then

$$\mathcal{M}_{v\parallel u} = \{ ab|d, ab|cd, ac|d, ac|bd, bd|a, bd|ac, cd|a, cd|ab, ad|bc \},$$

where the first 8 statements are there because of the first convention saying that  $a \perp\!\!\!\perp b \mid C [\mathcal{M}]$  whenever  $ab \cap C \neq \emptyset$  and the last statement  $ad|bc$  is added because of the second convention  $u \perp\!\!\!\perp u \mid C [\mathcal{M}]$  if and only if  $u \perp\!\!\!\perp (L \setminus C) \mid C [\mathcal{M}]$ . In case of  $\mathcal{R}$  one gets  $\mathcal{R}_{v\parallel u} = \mathcal{M}_{v\parallel u} \cup \{ ab|c, bd|c, ad|c \}$ , where  $bd|c$  is mapped by  $\rho$  to  $ab|c$  and  $ad|c$  is added because of the second convention.

*Remark on terminology.* Operations with supermodular set functions, which are analogous to those in Definitions 4.8 and 4.13, were discussed in [36, §7.1 and §10.2]. Moreover, the concept of tight replication of CI models is a counterpart of an analogous notion of *parallel extension* of matroids [31, §7.2] and of *polymatroids* [29, §2]. Nevertheless, the parallel extension of polymatroids has two slightly different reflections in the context of induced CI models. In the case that the considered polymatroid  $h$  is *tight* [9, §1.3] at the variable  $u \in N$ , which means that  $h(N) = h(N \setminus u)$ , it leads to our concept from Definition 4.13. If, however,  $h$  is not tight at  $u$  then it leads to the concept of parallel extension of CI models

introduced in 2004 by Matúš [28, end of § 4]. The extension we need for our purposes is the one that corresponds to a tight polymatroid. Therefore, to avoid a terminological clash, we have chosen a different term than parallel extension. Our terminology also reflects the crucial role of tightness.

**Lemma 4.15.** *The CI frames  $\text{st}$  and  $\text{sg}$  are closed under tight replication.*

**Proof.** Consider variable sets  $N, M$  assumed in the definition of tight replication and keep the corresponding notational conventions from Definition 4.13. Assume that  $\mathcal{M} \in \text{st}(N)$  is induced by a supermodular function  $r : \mathcal{P}(N) \rightarrow \mathbb{R}$ . Slight modification  $\hat{r}$  of  $r$  defined by

$$\hat{r}(S) := \begin{cases} r(S) - r(N) + r(L) & \text{if } u \in S, \\ r(S) & \text{if } u \notin S, \end{cases} \quad \text{for } S \subseteq N,$$

is another supermodular function on  $\mathcal{P}(N)$  which also induces  $\mathcal{M}$  but, moreover, satisfies  $\hat{r}(N) = \hat{r}(L)$ . The next step is to define

$$m(S) := \hat{r}(S \cap N) \cup \rho(S \cap M) \quad \text{for } S \subseteq NM.$$

For any  $ij|K \in \mathcal{K}(NM)$  one has  $\Delta m(ij|K) = \Delta \hat{r}(\rho(i), \rho(j)|\rho(K))$ . The latter expression is non-negative in case  $\rho(ij)|\rho(K) \in \mathcal{K}(N)$  by supermodularity of  $\hat{r}$ . Moreover, the expression vanishes if  $\rho(ij) \cap \rho(K) \neq \emptyset$ . In the remaining case  $\rho(i) = u = \rho(j)$  and  $C := \rho(K) \subseteq L$  one has

$$\Delta \hat{r}(\rho(i), \rho(j)|\rho(K)) = \Delta \hat{r}(u, u|C) = \hat{r}(C) - \hat{r}(uC) = \Delta \hat{r}(u, L \setminus C|C) \geq 0,$$

by supermodularity of  $\hat{r}$ . Hence,  $m$  is supermodular and it follows from the definition of tight replication that  $\mathcal{M}^m = \mathcal{M}_{v||u}$ .

Given  $\mathcal{M} \in \text{sg}(N)$ , it is enough to show that  $\mathcal{E} := \mathcal{M}_{v||u}$  is a semi-graphoid. Consider pairwise disjoint sets  $i, j, \ell, K \subseteq NM$  and the goal is to verify

$$ij|K, i\ell|jK \in \mathcal{E} \Leftrightarrow i\ell|K, ij|\ell K \in \mathcal{E}.$$

One can distinguish several cases determined by the mutual position of variables  $u$  and  $v$  relative to  $i, j, \ell, K$ . If  $uv \setminus ij\ell K \neq \emptyset$  then the conclusion follows from the facts that  $\mathcal{E}^{\downarrow N} = \mathcal{M}$  and  $\mathcal{E}^{\downarrow M}$  is a copy of  $\mathcal{M}$ . The same principle applies if  $u, v \in K$ . In cases  $|uv \cap K| = 1 = |uv \cap ij\ell|$  the first convention in the definition of  $\mathcal{M}_{v||u}$  is applied. This also extends to the case  $uv = j\ell$ . In the cases  $uv = ij$  and  $uv = i\ell$  one additionally applies the second convention in the definition of  $\mathcal{M}_{v||u}$  together with the assumption that  $\mathcal{M}$  is a semi-graphoid.  $\square$

**Remark.** Note in connection with Lemma 4.15 that we have not tried to answer the question whether  $\text{pr}$  is closed under tight replication. If it is the case then this certainly requires a technically complicated proof, but the point is that such a result is not needed for our purpose. Indeed, we know, by Lemma 4.7, that  $\text{pr}$  is self-adhesive, and, thus, self-adhesive at co-singletons.

**Lemma 4.16.** *Let  $\mathfrak{f}$  be a CI frame which is closed under copying, marginalization, and tight replication. Then, for any variable set  $N$  of cardinality  $n \geq 1$ , any model  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at every co-singleton.*

**Proof.** Consider  $\mathcal{M} \in \mathfrak{f}(N)$ , a subset  $L \subseteq N$  with  $|L| = n - 1$  (= a co-singleton), and a bijection  $\psi : N \rightarrow M$  such that  $N \cap M = L$  and  $\psi|_L = \text{id}_L$ . Put  $u := N \setminus L$  and  $v := M \setminus L$ . By the assumption,  $\mathcal{E} := \mathcal{M}_{v||u}$  belongs to  $\mathfrak{f}(NM)$ . The definition of tight replication allows one to observe that  $\mathcal{E}^{\downarrow N} = \mathcal{M}$ ,  $\mathcal{E}^{\downarrow M} = \psi(\mathcal{M})$ , and  $u \perp\!\!\!\perp v \mid L [\mathcal{E}]$ . Hence,  $\mathcal{E}$  is an adhesion of  $\mathcal{M}$  and  $\psi(\mathcal{M})$ .  $\square$

We now apply the observations from Sections 4.2 and 4.3.

**Corollary 4.17.** *If a CI frame  $\mathfrak{f}$  is closed under copying, marginalization, intersection, lifting, and tight replication and  $|N| \leq 3$  then  $\mathfrak{f}^\circ(N) = \mathfrak{f}(N)$ . In particular,  $\text{sg}^\circ(N) = \text{sg}(N)$  and  $\text{st}^\circ(N) = \text{st}(N)$  if  $|N| \leq 3$ .*

**Proof.** Combine Corollary 4.12, Lemma 4.16 and the fact that any  $\mathcal{M} \in \mathfrak{f}(N)$  is self-adhesive at  $L := N$ . Use Lemmas 3.1, 3.4, 4.10, and 4.15 for the second part.  $\square$

The second part of Corollary 4.17 is not surprising in light of Lemma 4.7 since  $\text{pr}(N) = \text{st}(N) = \text{sg}(N)$  if  $|N| \leq 3$ . On the other hand, Corollary 4.12 and Lemma 4.16 allow one to simplify testing self-adhesive models in case of large variable sets  $N$  with  $|N| \geq 4$ .

## 5. Self-adhesion operator and its iterating

In this section we discuss a self-adhesion operator which assigns a subframe  $\mathfrak{f}^\circ$  to an abstract CI frame  $\mathfrak{f}$ , closed under copying and marginalization. We prove the consistency results for such a construction, present an algorithm to compute the respective  $\mathfrak{f}^\circ$ -closure, and discuss the iterations of the self-adhesion operator.

Observe that copying commutes with marginalization: given  $\mathcal{M} \in \mathcal{K}(N)$ , a set  $L \subseteq N$ , and a bijection  $\varphi : N \rightarrow M$ , one has  $\varphi(\mathcal{M})^{\downarrow\varphi(L)} = \varphi_{|L}(\mathcal{M}^{\downarrow L})$ . Analogously, intersection commutes both with copying and marginalization: given  $\mathcal{M}, \mathcal{R} \in \mathcal{K}(N)$ , a bijection  $\varphi : N \rightarrow M$  and a set  $L \subseteq N$ , one has  $\varphi(\mathcal{M} \cap \mathcal{R}) = \varphi(\mathcal{M}) \cap \varphi(\mathcal{R})$  and  $(\mathcal{M} \cap \mathcal{R})^{\downarrow L} = \mathcal{M}^{\downarrow L} \cap \mathcal{R}^{\downarrow L}$ .

**Lemma 5.1.** *Let  $\mathfrak{f}$  be a CI frame closed under copying and marginalization.*

*Then  $N \mapsto \mathfrak{f}^\circ(N)$  defines a CI frame  $\mathfrak{f}^\circ$  which is also closed under copying and marginalization.*

- (i) *If  $\mathfrak{f}$  is, additionally, closed under intersection, then the same holds for  $\mathfrak{f}^\circ$ .*
- (ii) *If  $\mathfrak{f}$  is, additionally, closed under lifting, then the same holds for  $\mathfrak{f}^\circ$ .*
- (iii) *If  $\mathfrak{f}$  is, additionally, closed under intersection, lifting, and ascetic extension, then so is  $\mathfrak{f}^\circ$ .*

**Proof.** It is easy to see that  $\mathcal{K}(N) \in \mathfrak{f}^\circ(N)$  because  $\mathcal{K}(O) \in \mathfrak{f}(O)$  for any  $O \supseteq N$ . Hence,  $N \mapsto \mathfrak{f}^\circ(N)$  is an abstract CI frame in sense of Definition 2.1.

Consider  $\mathcal{M} \in \mathfrak{f}^\circ(N)$  and a bijection  $\varphi : N \rightarrow M$ . The goal is to show that  $\mathcal{R} := \varphi(\mathcal{M}) \in \mathfrak{f}^\circ(M)$ . Take  $L \subseteq M$  and choose a bijection  $\psi : M \rightarrow O$  such that  $L = M \cap O$ ,  $\psi_{|L} = \text{id}_L$ , and  $(O \setminus L) \cap N = \emptyset$ . By the remark n. 1 below Definition 4.3, it is enough to show that  $\mathcal{R}$  and  $\psi(\mathcal{R})$  have an adhesion in  $\mathfrak{f}(MO)$ . Put  $K := \varphi_{-1}(L) \cup (O \setminus L)$  and define a bijection  $\rho : N \rightarrow K$  by  $\rho(i) := i$  for  $i \in \varphi_{-1}(L)$  and  $\rho(i) := \psi(\varphi(i))$  for  $i \in N \setminus \varphi_{-1}(L)$ . Since  $\mathcal{M}$  is self-adhesive at  $\varphi_{-1}(L)$ , an extended model  $\mathcal{E} \in \mathfrak{f}(NK)$  exists with  $\mathcal{E}^{\downarrow N} = \mathcal{M}$ ,  $\mathcal{E}^{\downarrow K} = \rho(\mathcal{M})$  and  $(N \setminus \varphi_{-1}(L)) \perp\!\!\!\perp (K \setminus \varphi_{-1}(L)) \mid \varphi_{-1}(L) \mid \mathcal{E}$ . Extend  $\varphi$  by putting  $\varphi(i) := i$  for  $i \in K \setminus N = O \setminus L = O \setminus M$  and have  $\varphi : NK \rightarrow MO$ . Then  $\varphi(\mathcal{E}) \in \mathfrak{f}(MO)$  by the assumption on  $\mathfrak{f}$ . The fact that copying commutes with marginalization together with the identity  $\varphi_{|K} \circ \rho = \psi \circ \varphi_{|N}$  allows one to observe  $\varphi(\mathcal{E})^{\downarrow M} = \mathcal{R}$ ,  $\varphi(\mathcal{E})^{\downarrow O} = \psi(\mathcal{R})$  and  $(M \setminus L) \perp\!\!\!\perp (O \setminus L) \mid L \mid \varphi(\mathcal{E})$ .

Consider  $\mathcal{M} \in \mathfrak{f}^\circ(N)$  and  $M \subseteq N$ . To show  $\mathcal{M}^{\downarrow M} \in \mathfrak{f}^\circ(M)$  take  $L \subseteq M$  and choose a bijection  $\psi : M \rightarrow O$  with  $L = N \cap O$  and  $\psi_{|L} = \text{id}_L$ . Again, it is enough to show that  $\mathcal{M}^{\downarrow M}$  and  $\psi(\mathcal{M}^{\downarrow M})$  have an adhesion in  $\mathfrak{f}(MO)$ . Choose a set  $K$  with  $K \cap NO = \emptyset$  and extend  $\psi$  to a bijection from  $N$  to  $OK$ . Since  $\mathcal{M}$  is self-adhesive at  $L$ , a model  $\mathcal{E} \in \mathfrak{f}(NOK)$  exists with  $\mathcal{E}^{\downarrow N} = \mathcal{M}$ ,  $\mathcal{E}^{\downarrow OK} = \psi(\mathcal{M})$  and  $(N \setminus L) \perp\!\!\!\perp (OK \setminus L) \mid L \mid \mathcal{E}$ . The assumption on  $\mathfrak{f}$  gives  $\mathcal{E}^{\downarrow MO} \in \mathfrak{f}(MO)$ . Clearly,  $(\mathcal{E}^{\downarrow MO})^{\downarrow M} = \mathcal{E}^{\downarrow M} = (\mathcal{E}^{\downarrow N})^{\downarrow M} = \mathcal{M}^{\downarrow M}$ , and the fact that copying commutes with marginalization allows one to observe  $(\mathcal{E}^{\downarrow MO})^{\downarrow O} = \psi(\mathcal{M}^{\downarrow M})$ . It is immediate that  $(M \setminus L) \perp\!\!\!\perp (O \setminus L) \mid L \mid \mathcal{E}^{\downarrow MO}$ .

To verify (i) assume  $\mathcal{M}, \mathcal{R} \in \mathfrak{f}^\circ(N)$  and  $\varphi : N \rightarrow M$  with  $L := N \cap M$  and  $\varphi_{|L} = \text{id}_L$ . To show that  $\mathcal{M} \cap \mathcal{R}$  is self-adhesive at  $L$  one finds the respective extensions  $\mathcal{M} \mapsto \mathcal{E}_{\mathcal{M}}, \mathcal{R} \mapsto \mathcal{E}_{\mathcal{R}}$ , both in  $\mathfrak{f}(NM)$  and puts  $\mathcal{E} := \mathcal{E}_{\mathcal{M}} \cap \mathcal{E}_{\mathcal{R}} \in \mathfrak{f}(NM)$ , by the assumption on  $\mathfrak{f}$ . Since the intersection commutes both with marginalization and copying one gets  $\mathcal{E}^{\downarrow N} = \mathcal{M} \cap \mathcal{R}$  and  $\mathcal{E}^{\downarrow M} = \varphi(\mathcal{M} \cap \mathcal{R})$ . The definition of  $\mathcal{E}$  also implies  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid L \mid \mathcal{E}$ .

To verify (ii) assume  $\mathcal{M} \in \mathfrak{f}^\circ(N)$  and, without loss of generality,  $O \supseteq N$  with  $O \setminus N$  being a singleton  $o$ . We distinguish two cases. The first one is to show that  $\mathcal{M}_{N \nearrow O}$  is self-adhesive at  $L \subseteq N$  relative to  $\mathfrak{f}$ . Consider a bijection  $\varphi : O \rightarrow \tilde{M}$  with  $L = O \cap \tilde{M}$  and  $\varphi_{|L} = \text{id}_L$ ; put  $M := \tilde{M} \setminus \varphi(o)$ . The assumption on  $\mathcal{M}$  implies the existence of  $\mathcal{E} \in \mathfrak{f}(NM)$  with  $\mathcal{E}^{\downarrow N} = \mathcal{M}$ ,  $\mathcal{E}^{\downarrow M} = \varphi(\mathcal{M})$  and  $(N \setminus M) \perp\!\!\!\perp (M \setminus N) \mid L \mid \mathcal{E}$ . Then  $O\tilde{M}$  is a disjoint union of  $NM$  and  $o\varphi(o)$  and we can put  $\mathcal{R} := \mathcal{E}_{NM \nearrow O\tilde{M}}$ , which belongs to  $\mathfrak{f}(O\tilde{M})$ , by the assumption on  $\mathfrak{f}$ . The definition of lifting and the fact that it commutes with copying allow one to observe  $\mathcal{R}^{\downarrow O} = \mathcal{M}_{N \nearrow O}$ ,  $\mathcal{R}^{\downarrow \tilde{M}} = \varphi(\mathcal{M}_{N \nearrow O})$  and  $(O \setminus \tilde{M}) \perp\!\!\!\perp (\tilde{M} \setminus O) \mid L \mid \mathcal{R}$ .

The second case is to show that  $\mathcal{M}_{N \nearrow O}$  is self-adhesive at  $Lo$  (for  $L \subseteq N$ ) relative to  $\mathfrak{f}$ . Consider a bijection  $\varphi : O \rightarrow \tilde{M}$  with  $Lo := O \cap \tilde{M}$  and  $\varphi$  identical on  $Lo$ ; put  $M := \tilde{M} \setminus o$ . Choose an adhesion  $\mathcal{E} \in \mathfrak{f}(NM)$  of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$  (at  $L$ ). As  $O\tilde{M} = NMo$  put  $\mathcal{R} := \mathcal{E}_{NM \nearrow O\tilde{M}} \in \mathfrak{f}(O\tilde{M})$ . Analogous arguments allow one to show that  $\mathcal{R}$  is an adhesion of  $\mathcal{M}_{N \nearrow O}$  and  $\varphi(\mathcal{M}_{N \nearrow O})$ .

The verification of (iii) means, owing to (i) and (ii), to show that, given  $N \subseteq M$ , one has  $\mathfrak{f}^\circ(N) \subseteq \mathfrak{f}^\circ(M)$ . Since, however, we already know that  $\mathfrak{f}^\circ$  is closed under extension and intersection, it reduces to showing that  $\mathcal{K}(N) \in \mathfrak{f}^\circ(M)$  (see the comments below Definition 2.4). Thus, we put  $\mathcal{M} := \mathcal{K}(N)$  and regard  $\mathcal{M}$  as a CI model over  $M$ . We know that  $\mathcal{M} \in \mathfrak{f}(M)$  by the assumption on  $\mathfrak{f}$ . The goal is to verify that  $\mathcal{M} \in \mathfrak{f}^\circ(M)$ . To this end consider  $L \subseteq M$  and a bijection  $\varphi : M \rightarrow O$  with  $L = M \cap O$  and  $\varphi_{|L} = \text{id}_L$ . Let us put  $\mathcal{E}_1 := \mathcal{M}_{M \nearrow MO}$  and note that  $\mathcal{E}_1 \in \mathfrak{f}(MO)$  by the assumption of that  $\mathfrak{f}$  is closed under lifting. We also have  $\mathcal{E}_1^{\downarrow M} = \mathcal{M}$  and  $(M \setminus O) \perp\!\!\!\perp (O \setminus M) \mid L \mid \mathcal{E}_1$ . A special property of  $\mathcal{M}$ , namely  $ij|K \in \mathcal{M}, \tilde{K} \subseteq K \Rightarrow ij|\tilde{K} \in \mathcal{M}$ , together with the definition of lifting, allows one to derive that  $\varphi(\mathcal{M}) \subseteq \mathcal{E}_1$ . Analogously, we put  $\mathcal{E}_2 := \varphi(\mathcal{M})_{O \nearrow MO}$  and observe  $\mathcal{E}_2^{\downarrow O} = \varphi(\mathcal{M})$ ,  $(M \setminus O) \perp\!\!\!\perp (O \setminus M) \mid L \mid \mathcal{E}_2$  and  $\mathcal{M} \subseteq \mathcal{E}_2$ . Since  $\mathfrak{f}$  is closed under intersection,  $\mathcal{E} := \mathcal{E}_1 \cap \mathcal{E}_2 \in \mathfrak{f}(MO)$  is an adhesion of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$ .  $\square$

Note that the duality mapping commutes with copying: given  $\mathcal{M} \subseteq \mathcal{K}(N)$  and a bijection  $\varphi : N \rightarrow M$ , one has  $\varphi(\mathcal{M})^* = \varphi(\mathcal{M}^*)$ . On the other hand, it does not commute with marginalization: take  $N := \{a, b, c\}$ ,  $M := \{a, b\}$ , and  $\mathcal{M} := \{ab|\emptyset\}$ , in which case  $(\mathcal{M}^{\downarrow M})^* = \mathcal{M}$  while  $(\mathcal{M}^*)^{\downarrow M} = \emptyset$ . Therefore, a natural question is whether, given a CI frame  $\mathfrak{f}$  closed under duality, the frame  $\mathfrak{f}^\circ$  is closed under duality too. The answer is no: both  $\text{sg}^\circ$  (Example 7.2) and  $\text{st}^\circ$  (Example 7.3) are not closed under duality. An analogous question concerning the operation of tight replication remains open.

**Open problem 1.** *We would like to know whether either  $\text{sg}$  or  $\text{st}$  is an example of a CI frame  $\mathfrak{f}$  closed under tight replication whose self-adhesive shrinkage  $\mathfrak{f}^\circ$  is not closed under tight replication.*

Thus, we know from the basic claim in Lemma 5.1 that, if  $\mathfrak{f}$  is a CI frame allowing to introduce self-adhesive models, then  $\mathfrak{f}^\circ$  is another such a frame.

**Lemma 5.2.** *The self-adhesivity operator  $\mathfrak{f} \mapsto \mathfrak{f}^\circ$ , applied to CI frames  $\mathfrak{f}$  closed under copying and marginalization, is*

- anti-extensive:  $\mathfrak{f}^\circ \subseteq \mathfrak{f}$ , and
- isotone:  $\mathfrak{f}_1 \subseteq \mathfrak{f}_2$  implies  $\mathfrak{f}_1^\circ \subseteq \mathfrak{f}_2^\circ$ .

**Proof.** This is easy to observe on basis of Definition 4.3.  $\square$

**Corollary 5.3.** *Let  $\mathfrak{f}$  be a CI frame closed under copying and marginalization such that  $\text{pr} \subseteq \mathfrak{f}$ . Then  $\mathfrak{f}^\circ$  has the same properties and  $\text{pr} \subseteq \mathfrak{f}^\circ \subseteq \mathfrak{f}$ .*

**Proof.** Use the basic claim in Lemma 5.1. Then write  $\text{pr} = \text{pr}^\circ \subseteq \mathfrak{f}^\circ \subseteq \mathfrak{f}$  by Lemmas 2.5, 4.7 and 5.2.  $\square$

Thus, the new CI frame  $\mathfrak{f}^\circ$  yields a tighter approximation of  $\text{pr}$  than the original frame  $\mathfrak{f}$ . If, moreover,  $\mathfrak{f}$  is closed under intersection, then, by Lemma 5.1(i),  $\mathfrak{f}^\circ$  is also closed under intersection. Therefore, by observations from Section 2.2, for any  $N$ , the family  $\mathfrak{f}^\circ(N)$  can be characterized in terms of CI implications, namely the elements of the respective canonical implicational basis (see Definition 2.14). Elements of the canonical basis break down into permutational equivalence classes and each of them is equivalent to a collections of so-called Horn clauses [21], which have singletons as consequents. By Corollary 5.3, each of these valid implications for  $\mathfrak{f}^\circ(N)$  (see Section 2.3)

- is a valid probabilistic CI implication, that is, valid implication for  $\text{pr}(N)$ , while
- any valid implication for  $\mathfrak{f}(N)$  remains a valid implication for  $\mathfrak{f}^\circ(N)$ .

If additional assumptions on  $\mathfrak{f}$  are accepted then, by utilizing the algorithm in Lemma 4.5 as a subroutine, we can compute the  $\mathfrak{f}^\circ$ -closure of any model (see Definition 2.20).

**Lemma 5.4.** *Let  $\mathfrak{f}$  be a CI frame closed under copying, marginalization, intersection, lifting, and ascetic extension. Then the  $\mathfrak{f}^\circ$ -closure of any given  $\mathcal{M} \subseteq \mathcal{K}(N)$  can be computed by the following procedure:*

1. start with  $\mathcal{M}_0 := \mathcal{M}$ ,
2. iteratively compute  $\mathcal{M}_{i+1} := \bigcup_{L \subseteq N} \mathfrak{f}^\circ(\mathcal{M}_i|L)$  for  $i = 0, 1, \dots$ ,
3. if  $\mathcal{M}_{i+1} = \mathcal{M}_i$ , then output  $\mathfrak{f}^\circ(\mathcal{M}) := \mathcal{M}_i$  and stop.

**Proof.** Note that, by Lemma 5.1(iii),  $\mathfrak{f}^\circ$  is also closed under intersection and ascetic extension and the concept of  $\mathfrak{f}^\circ$ -closure makes sense (see Definition 2.20). An ascending chain of CI models  $\mathcal{M}_0 \subseteq \dots \subseteq \mathcal{M}_i \subseteq \dots \subseteq \mathcal{K}(N)$  is obtained by the iteration, which must eventually stabilize, say at  $\mathcal{M}_k$ . Since  $\mathfrak{f}^\circ(\mathcal{M}) \in \mathfrak{f}(N)$  is self-adhesive at any  $L \subseteq N$  relative to  $\mathfrak{f}$ ,  $\mathcal{M}_i \subseteq \mathfrak{f}^\circ(\mathcal{M})$  implies, by Definition 4.3,  $\mathfrak{f}^\circ(\mathcal{M}_i|L) \subseteq \mathfrak{f}^\circ(\mathcal{M})$ . Thus, one can show by induction that  $\mathcal{M}_i \subseteq \mathfrak{f}^\circ(\mathcal{M})$  at any stage of the algorithm. Since  $\mathcal{M}_k$  is a fixed point of the map  $\mathcal{M} \mapsto \bigcup_{L \subseteq N} \mathfrak{f}^\circ(\mathcal{M}|L)$ , it is self-adhesive relative to  $\mathfrak{f}$  at any subset  $L \subseteq N$  (simultaneously), and, thus, it equals to the  $\mathfrak{f}^\circ$ -closure of  $\mathcal{M}$ .  $\square$

The computation of  $\mathcal{M}_{i+1}$  in step 2 proceeds via Lemma 4.5 and requires one  $\mathfrak{f}$ -closure computation over a ground set of size  $2|N| - |L|$  for each  $L \subseteq N$ . We cannot provide an upper bound on the number of iterations until  $\mathcal{M}_i$  stabilizes. Matúš [27] gave an example of a model for which the computation of its  $\text{sg}(N)$ -closure using a similar iterative procedure requires an exponential number of iterations in  $|N|$ . There is no reason to expect the algorithm in Lemma 5.4 to have better worst-case complexity.

We wish to know whether or not the self-adhesivity operator  $\mathfrak{f} \mapsto \mathfrak{f}^\circ$  is idempotent, which means  $\mathfrak{f}^{\circ\circ} = \mathfrak{f}^\circ$ .<sup>6</sup> Since  $\text{sg}$  is closed under intersection, we were able to use a refined version of the procedure from Lemma 5.4 to compute the iterated self-adhesive models in  $\text{sg}^{\circ\circ}(N)$  in case  $|N| = 4$ . The basic idea of the algorithm is described in Appendix B. The result of our long computations was the conclusion that  $\text{sg}^{\circ\circ}(N) = \text{sg}^\circ(N)$  if  $|N| = 4$ . An analogous question concerning the  $\text{st}$ -frame remains open.

**Open problem 2.** *We would like to know whether  $\mathcal{M} \in \text{st}^\circ(N) \setminus \text{st}^{\circ\circ}(N)$  exists in case  $|N| = 4$ .*

## 6. Details on computations

This section is devoted to the computational aspects of our paper: here we explain how to use SAT solvers to compute self-adhesions of suitable CI frames.

<sup>6</sup> By Lemma 5.2,  $\mathfrak{f} \mapsto \mathfrak{f}^\circ$  would then be interpretable as the “interior” operation, which is a kind of counterpart of the “closure” operation discussed in Section 2.2.



### 6.1. Summary from the theoretical part

We first remind the reader which of our theoretical results from previous sections are used to make the computations possible. We certainly utilized the algorithm to compute the canonical basis (Definition 2.14) described in Appendix A, which works for arbitrary Moore family  $\mathcal{F}$  of subsets of a finite set  $X$ . Our algorithms are, however, tailored to compute the frame-based closure  $\mathfrak{f}(\mathcal{M})$  of a CI-model  $\mathcal{M}$ , where  $\mathfrak{f}$  is an appropriate abstract CI frame, as introduced in Section 2.3.

In connection with our central concept of self-adhesivity, we proposed an algorithm to compute the self-adhesive closure  $\mathfrak{f}^\circ(\mathcal{M})$  of a CI model  $\mathcal{M}$  relative to the self-adhesive frame  $\mathfrak{f}^\circ$  (Lemma 5.1), which needs the implementation of the basic  $\mathfrak{f}$ -closure. The algorithm is described in Appendix B and its consistency follows from Lemma 5.4. This global algorithm utilizes as a subroutine the procedure from Lemma 4.5 to compute (conditioned) self-adhesive closure  $\mathfrak{f}^\circ(\mathcal{M}|L)$  of  $\mathcal{M} \subseteq \mathcal{K}(N)$  at a given subset of variables  $L \subseteq N$ .

### 6.2. The SAT problem

The Boolean *satisfiability problem* (SAT) deals with Boolean formulas. A Boolean formula  $\phi$  is an expression using Boolean variables  $V_i, i \in X$ , indexed by a finite set  $X$ , and the Boolean operations of negation  $\neg$ , conjunction  $\wedge$ , and disjunction  $\vee$ . A mapping  $a : X \rightarrow \{\text{true}, \text{false}\}$  is called an *assignment* and it allows one to evaluate a given formula  $\phi$  by replacing each variable  $V_i$  by the Boolean value  $a(i)$  and computing negations, conjunctions, and disjunctions, resulting in a Boolean value denoted by  $\phi(a)$ . If  $\phi(a)$  is *true*, then  $a$  is called a *satisfying assignment*. Denote the set of all satisfying assignments of  $\phi$  by  $\text{SAT}(\phi)$ .

Two Boolean formulas  $\phi$  and  $\phi'$  are *equivalent* if  $\text{SAT}(\phi) = \text{SAT}(\phi')$ . A fundamental result on Boolean formulas states that, for every formula  $\phi$ , there is an equivalent formula  $\phi'$  in *conjunctive normal form* (CNF). In this normal form,  $\phi'$  is written as a conjunction of subformulas, called *clauses*. Each clause, in turn, is a disjunction of *literals*; and each literal is either a variable or a negated variable.

The SAT problem asks to decide whether a Boolean formula given in CNF has a satisfying assignment. In this form, the problem is famously NP-complete; see [1, Section 2.3]. Variants of this problem include counting all satisfying assignments (referred to as #SAT) and enumerating all satisfying assignments (ALLSAT). Remarkably efficient software for SAT, #SAT and ALLSAT exists and is constantly improving thanks to regular international competitions indexed at <https://www.satcompetition.org/>.

### 6.3. Using SAT solvers for self-adhesion computation

The SAT-based modelling is a powerful tool in combinatorics. Let  $X$  be the index set for variables in a Boolean formula. An assignment  $a$  may be identified with the subset  $\{i \in X : a(i) = \text{true}\} \subseteq X$ . The set of all satisfying assignments is then an element of  $\mathcal{P}(\mathcal{P}(X))$ . Choosing  $X := \mathcal{K}(N)$ , assignments are CI models over  $N$  and the set of satisfying assignments potentially forms a CI family.

By recognizing that a CI implication  $\bigwedge_{i=1}^k p_i \Rightarrow q$  (where  $p_1, \dots, p_k, q \in \mathcal{K}(N)$ ) is nothing but a CNF clause  $q \vee \neg p_1 \vee \dots \vee \neg p_k$ , it becomes apparent that the definition of a semi-graphoid in Section 3.1 for fixed  $N$  is, in fact, a Boolean formula  $\phi_N$  in CNF with  $\text{SAT}(\phi_N) = \text{sg}(N)$ . The frame of structural semi-graphoids is not defined by means of inference rules but, for each finite  $N$ , the Guigues–Duquenne procedure from Appendix A may be used to derive such a defining Boolean formula in conjunctive normal form.

Fix a CI family  $\mathfrak{f}(N)$  over  $N$  which is closed under intersection and assume that  $\phi$  is a Boolean formula in CNF such that  $\text{SAT}(\phi) = \mathfrak{f}(N)$ . We now show how the closure operator of  $\mathfrak{f}(N)$  can be computed using SAT solvers. Let  $\mathcal{M} \subseteq \mathcal{K}(N)$  be any set of CI statements and we fix  $ab|C \notin \mathcal{M}$ . To test whether  $ab|C \in \mathfrak{f}(\mathcal{M})$ , it suffices to test whether the CNF formula

$$\phi \wedge \bigwedge_{ij|K \in \mathcal{M}} V_{ij|K} \wedge \neg V_{ab|C} \tag{2}$$

is satisfiable. To see this, note that the satisfying assignments are precisely those elements of  $\mathfrak{f}(N)$  which contain  $\mathcal{M}$  and not  $ab|C$ . If one such model exists, then  $\mathfrak{f}(\mathcal{M})$ , being the intersection of all models containing  $\mathcal{M}$ , does not contain  $ab|C$ ; otherwise the closure must contain  $ab|C$  as all supersets of  $\mathcal{M}$  in  $\mathfrak{f}(N)$  contain it.

It is known from [27] that even in the simple frame of semi-graphoids, deciding whether a CI statement belongs to the closure of a given model may require a number of steps exponential in the ground set size. Despite this, SAT solvers perform exceedingly well in practice as inference engines for CI implication. We employed the solver CaDiCaL [44], winner of the SAT Race 2019, which uses conflict-driven clause learning. This strategy makes it especially powerful in our setting, where the same formula has to be checked many times under different specializations of the variables (only varying the statement  $ab|C$  in (2)).

It is evident from the above description that the closure operator of  $\mathfrak{f}$  is at the core of all our computations related to the self-adhesion operator. We employ a state-of-the-art incremental SAT solver to make this computation as fast as possible; all other code is less critical and mainly organizes the high-level computation, determining which closures are needed and combining the results.



#### 6.4. Using LP solvers for self-adhesion computation

Self-adhesivity of polymatroids (see later [Definition 8.1](#)) has been used to derive further information inequalities (cf. Sections [4.1](#) and [8](#)). This process follows the same pattern as our algorithms, except that Boolean formulas are replaced by polyhedral cones (whose elements are polymatroids) and, consequently, SAT solvers must be replaced by *linear programming* (LP) solvers. Such computations have been carried out at a large scale in [\[4\]](#). The accompanying webpage

<https://mathrepo.mis.mpg.de/SelfadhesiveGaussianCI/>

goes into more details about the time and memory demands of SAT versus LP solvers. In the computation of self-adhesions of orientable gaussoids (using a SAT solver) versus self-adhesions of structural gaussoids (using an LP solver), the SAT solver was on average 5 times faster even though it had to deal with formulas with twice as many variables as the LP solver.

However, this comparison is unfair as the SAT solver was integrated into the process and designed to work incrementally, while the LP solver was an external program which had to be called repeatedly. About two thirds of the time it takes to compute the self-adhesive closure using the LP solver was reportedly spent on writing the linear program and starting the external LP solver. Compared to the abundance of excellent SAT solvers, it is more difficult to find a fast LP solver which is freely available software, easy to integrate into domain-specific large-scale computations and supports exact rational arithmetic which is necessary to guarantee the robustness of our results. Our computations also did not exploit recent advances in pre-processing of linear programs [\[19\]](#).

### 7. Computational results for small numbers of variables

In this section we present the results of our computations. Our implementation relies on the following software packages: the SCIP optimization suite [\[47\]](#) and its LP solver SoPlex [\[48\]](#); discrete geometry software normaliz [\[46\]](#); SAT solvers CaDiCaL [\[44\]](#), GANAK [\[49\]](#) and nbc\_minisat\_all [\[50\]](#); the Clnet framework [\[45\]](#); as well as Mathematica [\[51\]](#). More details are available on our supplementary website <https://cinet.link/data/selfadhe-lattices>. We also give a few instructive examples obtained as their results. A summary of numerical data about the lattices we consider in case  $|N| = 4$  is available in [Appendix C](#).

#### 7.1. Semi-graphoids over 4 variables

Let us start our overview by discussing the family of semi-graphoids  $\text{sg}(N)$  in case  $|N| = 4$ . One has 26 424 semi-graphoids over any 4-element variable set and they form 1512 permutational equivalence classes. The lattice  $(\text{sg}(N), \subseteq)$  has in this case 181 (meet) irreducible elements which fall into 20 permutational types. On the other hand, it only has 37 coatoms falling down into 10 types; thus, the lattice is not coatomistic. In particular, the description of  $(\text{sg}(N), \subseteq)$  in terms (of types) of irreducible models is relatively complicated.

On the other hand, the description of the lattice  $(\text{sg}(N), \subseteq)$  in terms of implicational generators is very easy. The canonical implicational basis  $\mathbb{G}_{\mathcal{F}}$  for  $\mathcal{F} = \text{sg}(N)$  in case  $|N| = 4$  consists of two permutational types only, namely

$$\{ij|\emptyset, il|j\} \rightarrow \{il|\emptyset, ij|l\} \quad \text{and} \quad \{ij|k, il|jk\} \rightarrow \{il|k, ij|kl\}.$$

Thus, the lattice  $(\text{sg}(N), \subseteq)$  is implicatively perfect. This is not a surprise because the semi-graphoidal frame has, by its definition, a finite “axiomatic” characterization. Informally, by such a characterization of a frame  $\mathfrak{f}$  we mean that there is a finite number of schemes which, for each  $|N|$ , yield an implicational generator for the  $\mathfrak{f}$ -closure operation.

#### 7.2. Structural semi-graphoids over 4 variables

The family  $\text{st}(N)$  in case  $|N| = 4$  contains 22 108 models, which fall down into 1285 permutational equivalence classes. The lattice  $(\text{st}(N), \subseteq)$  is known to be coatomistic for any  $N$ . Moreover, in the discussed case  $|N| = 4$ , it has the same 37 coatoms as the semi-graphoidal lattice  $(\text{sg}(N), \subseteq)$ . Thus, the description of the family  $\text{st}(N)$  in terms of irreducible models is much easier in this case than the description of  $\text{sg}(N)$  in these terms.

As concerns the description of the lattice  $(\text{st}(N), \subseteq)$  in terms of implicational generators, we computed the canonical implicational basis  $\mathbb{G}_{\mathcal{F}}$  for  $\mathcal{F} = \text{st}(N)$  in case  $|N| = 4$ . What we found out is that it decomposes into 7 permutational types. More specifically, it contains two semi-graphoidal types mentioned in [Section 7.1](#) and five additional permutational types listed below:

$$\begin{aligned} \text{(E:1)} \quad & \{ij|k, ik|l, il|j\} \rightarrow \{ij|l, ik|j, il|k\}, \\ \text{(E:2)} \quad & \{ij|k, il|j, jk|l, kl|i\} \rightarrow \{ij|l, il|k, jk|i, kl|j\}, \\ \text{(E:3)} \quad & \{ij|kl, ik|\emptyset, jl|\emptyset, kl|ij\} \rightarrow \{ij|\emptyset, ik|jl, jl|ik, kl|\emptyset\}, \\ \text{(E:4)} \quad & \{ij|\emptyset, ij|kl, kl|i, kl|j\} \rightarrow \{ij|k, ij|l, kl|\emptyset, kl|ij\}, \\ \text{(E:5)} \quad & \{ij|k, jk|i, il|j, kl|\emptyset\} \rightarrow \{ij|kl, jk|i, il|\emptyset, kl|j\}. \end{aligned}$$

Note that our results agree with the analysis presented in [\[37, §V.A\]](#), or with the rules A3–A7 from [\[34\]](#). In particular, the lattice  $(\text{st}(N), \subseteq)$  is implicatively perfect in case  $|N| = 4$ . Moreover, both the semi-graphoidal rules and the rules (E:1)–(E:5) can be generalized to any higher cardinality than  $|N| \geq 4$  [\[6,7\]](#).

### 7.3. Self-adhesive semi-graphoids over 4 variables

The family  $\text{sg}^\circ(N)$  in case  $|N| = 4$  contains 23 190 models, which fall down into 1352 permutational equivalence classes. The number of (meet) irreducible elements in this lattice  $(\text{sg}^\circ(N), \subseteq)$  is 385 and they break into 29 permutational types. Note that, in comparison with  $\text{sg}(N)$ , the lattice is smaller but the number of its irreducible elements is higher.

The lattice  $(\text{sg}^\circ(N), \subseteq)$  has 31 coatoms in this case, which break into 9 types. Therefore, it is not coatomistic. There is only one permutational type of semi-graphoidal coatoms which is not self-adhesive. It leads to an example of a CI model in  $\text{st}(N) \setminus \text{sg}^\circ(N)$ , the largest one in case  $|N| = 4$ .

**Example 7.1** ( $\text{st}(N) \setminus \text{sg}^\circ(N) \neq \emptyset$ ).

Consider  $N := \{a, b, c, d\}$  and put

$$\mathcal{M} := \{ ab|c, ab|d, ab|cd, ac|b, ad|b, bc|a, bd|a, cd|\emptyset, cd|a, cd|b \}.$$

This model  $\mathcal{M}$  is a coatom both in  $(\text{sg}(N), \subseteq)$  and  $(\text{st}(N), \subseteq)$  (which elements coincide in case  $|N| = 4$ ). Nevertheless,  $\mathcal{M} \notin \text{sg}^\circ(N)$ . To observe this we show below that  $\mathcal{M}$  is not self-adhesive at  $L := bd$  relative to  $\text{sg}$ . We apply Lemma 4.5 with  $f := \text{sg}$  for this purpose. Fix a bijection  $\varphi : N \rightarrow M := \{\bar{a}, b, \bar{c}, d\}$ , where  $\varphi(a) := \bar{a}$ ,  $\varphi(b) := b$ ,  $\varphi(c) := \bar{c}$ , and  $\varphi(d) := d$ . To show that  $bc|\emptyset \in \text{sg}^\circ(\mathcal{M}|L) \setminus \mathcal{M}$  it is enough to verify that  $bc|\emptyset \in \text{sg}^\circ(\mathcal{R}|L)$  for a subset  $\mathcal{R} := \{ ab|d, ad|b, bc|a, bd|a, cd|\emptyset \}$  of  $\mathcal{M}$ .

By Lemma 4.5, this means that  $bc|\emptyset$  belongs to the semi-graphoidal closure of the union of  $\mathcal{R}$ ,  $\varphi(\mathcal{R})$ , and  $[ac, \bar{a}\bar{c}|bd]$ . The following is one of possible semi-graphoidal derivation sequences:

1.  $ad|b \in \mathcal{R}, [a, \bar{a}|bd] \rightarrow ad|\bar{a}b \ \& \ bd|\bar{a} \in \varphi(\mathcal{R}) \rightarrow ad|\bar{a}$ ,
2.  $ab|d \in \mathcal{R}, [a, \bar{a}\bar{c}|bd] \rightarrow [a, b\bar{c}|\bar{a}d] \ \& \ ad|\bar{a} \text{ (1.)} \rightarrow ab|\bar{a}\bar{c}$ ,
3.  $\bar{a}d|b \in \varphi(\mathcal{R}), [\bar{a}, ac|bd] \rightarrow \bar{a}c|ab \ \& \ bc|a \in \mathcal{R} \rightarrow bc|\bar{a}\bar{a}$ ,
4.  $\bar{a}b|d \in \varphi(\mathcal{R}), [\bar{a}, c|bd] \rightarrow \bar{a}c|d \ \& \ cd|\emptyset \in \mathcal{R} \rightarrow \bar{a}c|\emptyset$ ,
5.  $b\bar{c}|\bar{a} \in \varphi(\mathcal{R}) \ \& \ ab|\bar{a}\bar{c} \text{ (2.)} \rightarrow ab|\bar{a} \ \& \ bc|\bar{a}\bar{a} \text{ (3.)} \rightarrow bc|\bar{a} \ \& \ \bar{a}c|\emptyset \text{ (4.)} \rightarrow bc|\emptyset$ .

Remark that the derivation like  $ab|d \ \& \ [a, \bar{a}\bar{c}|bd] \rightarrow [a, b\bar{c}|\bar{a}d]$  is based on the application of semi-graphoidal inference in its global mode (see Definition 2.3). In this case, this is a shorthand for two-stage derivation:

$$[a \perp\!\!\!\perp b \mid d \ \& \ a \perp\!\!\!\perp \bar{a}\bar{c} \mid bd] \rightarrow a \perp\!\!\!\perp b\bar{a}\bar{c} \mid d \rightarrow a \perp\!\!\!\perp b\bar{c} \mid \bar{a}d.$$

Note that the only set  $L$  at which the model  $\mathcal{M}$  is self-adhesive is  $cd$ .

Of course, the derivation sequence from Example 7.1 also confirms that the considered semi-graphoid  $\mathcal{R}$  is not self-adhesive at  $L = bd$ . This is a kind of a certificate (= a direct proof) of the validity of a later implicational rule (SI:3) for self-adhesive semi-graphoids; take  $i := b, j := a, k := d$ , and  $\ell := c$  in (SI:3).

The simplest example of a self-adhesive semi-graphoid that is not structural is given in Example 3.2. Its minor modification below shows that  $\text{sg}^\circ$  is not closed under duality, analogously to the  $\text{pr}$ -frame (see Example 2.6).

**Example 7.2** ( $\text{sg}^\circ(N) \setminus \text{st}(N) \neq \emptyset$ ,  $\text{sg}^\circ$  is not Closed Under Duality).

Put  $N := \{a, b, c, d\}$  and consider the CI model  $\mathcal{M} := \{ ab|c, ac|d, ad|b, bc|d \}$ . It appears to be a self-adhesive semi-graphoid which is not structural: apply the rule (E:1) presented in Section 7.2 with  $i := a, j := b, k := c$ , and  $\ell := d$ . Nevertheless, its dual model  $\mathcal{M}^* = \{ ab|d, ac|b, ad|c, bc|a \}$  is not self-adhesive. This follows from the below-mentioned property (SE:1) of self-adhesive semi-graphoids, where  $i := a, j := c, k := b$ , and  $\ell := d$ . Note that  $\mathcal{M}, \mathcal{M}^* \notin \text{st}(N)$  and, thus, they are not probabilistic CI models.

Note in this context that, if  $|N| = 4$  then 48 permutational types of models  $\mathcal{M} \in \text{sg}^\circ(N)$  with  $\mathcal{M}^* \notin \text{sg}^\circ(N)$  exist.

We also computed the canonical implicational basis  $\mathbb{G}_{\mathcal{F}}$  for  $\mathcal{F} = \text{sg}^\circ(N)$  in case  $|N| = 4$ . It decomposes into 9 permutational types: besides two semi-graphoidal types mentioned in Section 7.1 it has 7 additional types and we divided them into two subgroups. The implications in the first subgroup can be interpreted as weaker versions of probabilistic equivalences from [37, §V.A], recalled in Section 7.2, and, therefore, we use the label SE to denote them:

- (SE:1)  $\{ ij|k, ik|\ell, jk|i, i\ell|j \} \rightarrow \{ ij|\ell, ik|j, i\ell|k \}$ ,  
 (SE:2)  $\{ ij|k, ij|k\ell, ik|j, jk|i, jk|i\ell, i\ell|\emptyset, i\ell|j, k\ell|\emptyset \} \rightarrow \{ k\ell|j \}$ .

Indeed, the implication (SE:1) weakens the property (E:1) with an additional antecedent  $jk|i$ , while the rule (SE:2) weakens the rule (E:5): the 4 additional antecedents are  $ij|k\ell, ik|j, jk|i$ , and  $i\ell|\emptyset$ .

The implications in the second subgroup can be interpreted as weaker versions of probabilistic implications presented in [37, §V.B], for which reason we use the label SI to denote them:

- (SI:1)  $\{ ij|k, ik|j, jk|i, jk|\ell, i\ell|j \} \rightarrow \{ ij|\ell \}$ ,  
 (SI:2)  $\{ ij|k, ik|j, jk|i, jk|i\ell, i\ell|j \} \rightarrow \{ ij|k\ell \}$ ,  
 (SI:3)  $\{ ij|k, ik|j, jk|i, i\ell|j, k\ell|\emptyset \} \rightarrow \{ i\ell|\emptyset \}$ ,

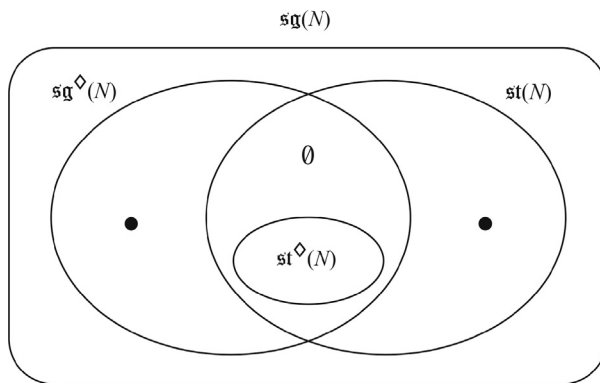


Fig. 4. Relation between considered families of CI models in case  $|N| = 4$ .

(SI:4)  $\{ij|k, ik|j, jk|i, i\ell|j, k\ell|i\} \rightarrow \{i\ell|k\}$ ,  
 (SI:5)  $\{ij|k, ij|\ell, ik|j, jk|i, k\ell|\emptyset\} \rightarrow \{ij|\emptyset, ik|\emptyset, jk|\emptyset\}$ .

This is a comparison with the implications from [37, §V.B]:

- remove the antecedent  $jk|i$  in (SI:1) to get (I:10) or  $ik|j$  to get (I:11),
- remove  $jk|i$  in (SI:2) to get (I:17) or  $ik|j$  to get (I:18),
- remove  $ik|j$  in (SI:3) to get (I:5) or  $jk|i$  to get (I:6),
- remove  $ik|j$  in (SI:4) to get (I:8) or  $jk|i$  to get (I:12),
- remove  $jk|i$  and  $jk|\emptyset$  in (SI:5) to get (I:2) or  $ij|k$  and  $ij|\emptyset$  to get (I:4).

The result of our analysis was also the observation that the lattice  $(sg(N), \subseteq)$  is **not** implicatively perfect in case  $|N| = 4$ . Indeed, the implication (SE:2) belongs to its canonical implicational basis but it is not perfect: its antecedent set is the union of CI statements from rules (SI:2) and (SI:3).

As illustrated by Example 2.15, the rule (SE:2) can further be simplified:

(SE:2)  $\{ij|k, ik|j, jk|i, jk|i\ell, i\ell|j, k\ell|\emptyset\} \rightarrow \{k\ell|j\}$ .

We also confirmed independently the validity of the rules (SE:1)–(SE:2) and (SI:1)–(SI:5) within the frame  $sg^\diamond$  by finding their certificates in the form of semi-graphoidal derivation sequences, as illustrated in Example 7.1. Specifically, we found out that the antecedent set in (SE:1) is not self-adhesive at sets  $ij$  and  $jk$ , while the antecedent sets in (SE:2) and in the rules labelled by SI are not self-adhesive at the sets  $ij$ ,  $ik$ , and  $jk$ .

#### 7.4. Self-adhesive structural semi-graphoids: 4 variables

The family  $st^\diamond(N)$  in case  $|N| = 4$  contains 20 968 models, falling down into 1224 permutational equivalence classes. The number of (meet) irreducible elements in the lattice  $(st^\diamond(N), \subseteq)$  is 85 and they break into 13 permutational types, presented in the form of a concise catalogue in Appendix D.

The same phenomenon was observed as in Section 7.3: in comparison with  $(st(N), \subseteq)$ , the number of elements of the lattice decreases while the number of its irreducible elements increases. The set of coatoms of  $(st^\diamond(N), \subseteq)$  is the same as in case of  $(sg^\diamond(N), \subseteq)$ , namely 31 of them falling into 9 permutational types. In particular, the lattice  $(st^\diamond(N), \subseteq)$  is not coatomistic for  $|N| = 4$ .

We also computed the canonical implicational basis  $\mathbb{G}_{\mathcal{F}}$  for  $\mathcal{F} = st^\diamond(N)$  in case  $|N| = 4$ . It decomposes into 12 permutational types: it consists of two semi-graphoidal types mentioned in Section 7.1, five types (E:1)–(E:5) from Section 7.2 and five types (SI:1)–(SI:5) presented in Section 7.3. It makes no problem to observe that all these implications are perfect, implying that the lattice  $(st^\diamond(N), \subseteq)$  is implicatively perfect in case  $|N| = 4$ .

Our computation also revealed 30 types of models  $\mathcal{M} \in st^\diamond(N)$  such that  $\mathcal{M}^* \notin st^\diamond(N)$ . One of these examples is as follows.

#### Example 7.3 ( $st^\diamond$ is not Closed Under Duality).

Put  $N := \{a, b, c, d\}$ . The model  $\mathcal{M} := \{ab|d, ac|d, ad|b, bc|\emptyset, bc|d\}$  belongs to  $st^\diamond(N)$  while its dual model not:  $\mathcal{M}^* := \{ab|c, ac|b, ad|c, bc|a, bc|ad\} \notin st^\diamond(N)$ . The latter fact follows from the rule (SI:2) above where  $i := a, j := c, k := b$ , and  $\ell := d$ .

A kind of surprise for us was the observation that  $sg^\diamond(N) \cap st(N) = st^\diamond(N)$  if  $|N| = 4$  (see Fig. 4 for illustration); thus, self-adhesivity relative to  $sg$  and relative to  $st$  yield the same restriction. Therefore, we have a natural question.

**Open problem 3.** We would like to know whether the relation  $\text{sg}^\circ(N) \cap \text{st}(N) = \text{st}^\circ(N)$  holds in general.

7.5. Graphoids, compositional semi-graphoids, and compositional graphoids: 4 variables

A result of our computations was the finding that, in case  $|N| = 4$ , the family  $\text{gr}(N)$  contains 6482 elements, which fall down into 421 permutational types. The number of (meet) irreducible models in  $\text{gr}(N)$  is 408 and they decompose into 32 types. The lattice  $(\text{gr}(N), \subseteq)$  has 14 coatoms of 4 types; thus, it is not coatomistic. The family  $\text{gr}^\circ(N)$  in case  $|N| = 4$  coincides with  $\text{gr}(N)$ . Note that we also found that, in case  $|N| = 4$ , precisely 369 permutational types of graphoids are probabilistically representable.

Since, for any  $N$ ,  $(\text{gr}^*(N), \subseteq)$  is order-isomorphic to  $(\text{gr}(N), \subseteq)$ , the numbers of elements, irreducible elements, and coatoms of  $(\text{gr}^*(N), \subseteq)$  are the same as in case of graphoids. What is, however, different is that  $\text{gr}^*(N) \neq (\text{gr}^*)^\circ(N)$  in case  $|N| = 4$ . The lattice  $(\text{gr}^*)^\circ(N)$  has only 6182 (instead of 6482) elements in 404 permutational types. It has 14 coatoms in 4 types and 428 irreducible elements falling to 33 types. This is the first concrete example of a frame  $\mathfrak{f}$  such that both  $\mathfrak{f}$  and  $\mathfrak{f}^*$  admit self-adhesion and  $(\mathfrak{f}^*)^\circ \neq (\mathfrak{f}^\circ)^*$ , i.e., the operations of duality and self-adhesion on a CI frame do not commute in general.

The family  $\text{cgr}(N) = \text{gr}(N) \cap \text{gr}^*(N)$  has 2084 elements in case  $|N| = 4$  which fall into 157 types. It has 470 irreducible models of 30 types and 6 coatoms of 1 type. In particular, the lattice is not coatomistic. Of course,  $\text{cgr}(N) \subseteq \text{gr}(N) \subseteq \text{sg}^\circ(N)$  in this case. Like  $\text{gr}(N)$  in case  $|N| = 4$ ,  $\text{cgr}(N) = \text{cgr}^\circ(N)$ .

7.6. The case of 5 variables

We had the following special goal in this case. There are 1319 permutational types of coatoms of the lattice  $(\text{st}(N), \subseteq)$  of structural semi-graphoids over 5 variables. Note that the overall number of coatoms of this lattice is 117 978.

In general, the coatoms of this lattice correspond to the extreme rays of the cone of standardized supermodular functions, that is, to supermodular functions  $m : \mathcal{P}(N) \rightarrow \mathbb{R}$  satisfying  $m(K) = 0$  if  $K \subseteq N$ ,  $|K| \leq 1$ , see [39, Corollary 30]. Thus, the coatoms for  $|N| = 5$  were obtained by computing the extreme rays of that cone [38]. We raised a couple of questions:

- Which of the permutational types belong to  $\text{sg}^\circ(N)$ ?
- Which of them belong to  $\text{st}^\circ(N)$ ?

Surprising answers were that only 154 of those types belong to  $\text{sg}^\circ(N)$  and these coincide with the types belonging to  $\text{st}^\circ(N)$ . This observation seems to support the conjecture indicated in Open problem 3. Consequences of those results are discussed in detail in Section 8.

**8. Application to entropic region delimitation**

Given a discrete random vector  $\xi = [\xi_i]_{i \in N}$  over a variable set  $N$  and  $I \subseteq N$ , the (Shannon’s) *entropy* (of its sub-vector) for  $I$  is the number

$$h_\xi(I) := \sum_{x \in X_N : p_N(x) > 0} p_N(x) \cdot \log \left( \frac{1}{p_I(x)} \right) \quad (\text{see Section 2.1}).$$

Well-known facts are that the *entropy function*  $h_\xi : \mathcal{P}(N) \rightarrow \mathbb{R}$  is submodular, non-decreasing with respect to inclusion, and satisfies  $h_\xi(\emptyset) = 0$ . Thus, in terms of Section 3.2, it is (the rank function of) a polymatroid over  $N$ .

A polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$  is then called *entropic* if  $h = h_\xi$  for some discrete random vector  $\xi$  over  $N$ . It is called *almost entropic* if it is a limit of entropic polymatroids (in the Euclidean topology on  $\mathbb{R}^{\mathcal{P}(N)}$ ). The *entropic region*, also named the *entropy region*, is the set of all entropic polymatroids and the *almost entropic region* is the set of all almost entropic polymatroids; the latter set is known to be a closed cone in  $\mathbb{R}^{\mathcal{P}(N)}$  [41, § 15.1].

*Information inequalities*, that is, linear inequalities valid for vectors in the (almost) entropic region, are one of the research topics in information theory. In this paper we discuss *unconstrained* information inequalities, valid for almost entropic vectors. Besides classic Shannon’s inequalities, defining the cone of polymatroids, a number of independent such inequalities has been found; the first of them in 1998 [43]. The techniques to derive these inequalities follow the scheme of the so-called *copy lemma* [13] and inspired the concept of a self-adhesive polymatroid by Matúš [29], recalled below in analogy with Definition 4.3.

**Definition 8.1 (Self-Adhesive Polymatroid).**

Given a bijection  $\varphi : N \rightarrow M$ , the  $\varphi$ -*copy* of a polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$  is the polymatroid  $h \circ \varphi_{-1} \in \mathbb{R}^{\mathcal{P}(M)}$ . A pair of sets  $N, M$  is called *modular* in a polymatroid  $\hat{h} \in \mathbb{R}^{\mathcal{P}(NM)}$  if  $\hat{h}(N) + \hat{h}(M) = \hat{h}(NM) + \hat{h}(N \cap M)$ . A polymatroid  $\hat{h} \in \mathbb{R}^{\mathcal{P}(NM)}$  extends  $h \in \mathbb{R}^{\mathcal{P}(N)}$  if  $h$  is the restriction of  $\hat{h}$  to  $\mathcal{P}(N)$ . We say that a polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$  is *self-adhesive at a set*  $L \subseteq N$

if,

- $\forall \varphi : N \rightarrow M$  bijection with  $L = N \cap M$  and  $\varphi|_L = \text{id}_L$
- $\exists$  a polymatroid  $\hat{h} \in \mathbb{R}^{\mathcal{P}(NM)}$  extending both  $h$  and its  $\varphi$ -copy such that  $N, M$  is a modular pair in  $\hat{h}$ .

A polymatroid over  $N$  is *self-adhesive* if it is self-adhesive at every  $L \subseteq N$ .

Recall from Section 3.2 that a polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$  induces a CI model  $\mathcal{M} \subseteq \mathcal{K}(N)$  as follows:  $ij|K \in \mathcal{M}$  if and only if  $\Delta h(ij|K) = 0$ .

**Lemma 8.2.** *The following statements are true.*

- (a) *The set of self-adhesive polymatroids is a polyhedral cone in  $\mathbb{R}^{\mathcal{P}(N)}$ .*
- (b) *Every almost entropic polymatroid is self-adhesive.*
- (c) *Every self-adhesive polymatroid  $h$  over  $N$  induces a CI model in  $\text{st}^\circ(N)$ .*

**Proof.** To show (a) is true, consider a set  $L \subseteq N$  and the respective bijection  $\varphi : N \rightarrow M$ . Realize that the set of polymatroids  $\hat{h} \in \mathbb{R}^{\mathcal{P}(NM)}$  in which  $N, M$  is a modular pair and that satisfy  $\hat{h}(A) = h(\varphi(A))$  for  $A \subseteq N$  is a polyhedral cone. A polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$  is self-adhesive at  $L$  if and only if it belongs to the projection of that set of vectors  $\hat{h}$  from  $\mathbb{R}^{\mathcal{P}(NM)}$  to  $\mathbb{R}^{\mathcal{P}(N)}$ . Basic facts from polyhedral geometry are that the class of polyhedral cones is closed under projections and finite intersections. The application of these facts yields (a).

For (b) we first realize that any entropic polymatroid is self-adhesive. The arguments for this from [29, Corollary 1] are analogous to those in our proof of Lemma 4.7. Then (a) implies that set of self-adhesive polymatroids over  $N$  is closed, which forces that it contains the closure of the entropic region.

As concerns (c), recall from the remark n. 1 in the end of Section 3.2 that  $\mathcal{M} \in \text{st}(N)$  if and only if it is induced by a polymatroid over  $N$ . Given  $\mathcal{M} \subseteq \mathcal{K}(N)$  induced by a self-adhesive polymatroid  $h \in \mathbb{R}^{\mathcal{P}(N)}$ , one needs to observe that  $\mathcal{M}$  is self-adhesive at (any)  $L \subseteq N$  relative to  $\text{st}$ . Having fixed the corresponding bijection  $\varphi : N \rightarrow M$ , take the polymatroid  $\hat{h} \in \mathbb{R}^{\mathcal{P}(NM)}$  from Definition 8.1 and it is straightforward to verify that it induces  $\hat{\mathcal{M}} \in \text{st}(NM)$  which is an adhesive extension of  $\mathcal{M}$  and  $\varphi(\mathcal{M})$ .  $\square$

Lemma 8.2 has the following consequence.

**Corollary 8.3.** *If a polymatroid  $h$  over  $N$  induces a CI model  $\mathcal{M}$  outside  $\text{st}^\circ(N)$  then there is a valid (unconstrained) information inequality which is not true for  $h$  and none of positive multiples of  $h$  is almost entropic.*

**Proof.** Realize that any facet-defining inequality for the cone of self-adhesive polymatroids is a valid information inequality.  $\square$

To interpret the results from Section 7.6 we recall some geometric facts. A polymatroid  $h$  over  $N$ ,  $|N| \geq 1$ , is called *tight* if  $h(N) = h(N \setminus i)$  for any  $i \in N$  and *modular* if  $\Delta h(ij|K) = 0$  for any  $ij|K \in \mathcal{K}(N)$ . Every polymatroid  $h$  then decomposes into its modular part and its tight part  $\tilde{h}$  inducing the same CI model as  $h$ . Moreover, there exists a linear one-to-one mapping between tight polymatroids and standardized supermodular functions which preserves the induced CI model, see [39, formulas (32) in §7.2]. This allows one to observe, using Corollary 30 in [39, Appendix B], that a polymatroid  $h$  over  $N$  generates a non-modular (= tight) extreme ray of the polymatroidal cone if and only if it induces a coatom in the lattice  $(\text{st}(N), \subseteq)$ .

Thus, in light of Corollary 8.3, the computational results from Section 7.6 imply that most of the 117 978 extreme rays of the cone of tight polymatroids over a 5-element set are outside the almost entropic region.

Being motivated by this perspective, we put (for every variable set  $N$ ):

$$\overline{\text{pr}}(N) := \{ \mathcal{M} \subseteq \mathcal{K}(N) : \mathcal{M} \text{ is induced by an almost entropic polymatroid over } N \},$$

which defines the *almost entropic* CI frame  $\overline{\text{pr}}$ . It seems to coincide with the “approximate probabilistic frame” suggested by one of the reviewers: one has  $\mathcal{M} \in \overline{\text{pr}}(N)$  if  $\exists \delta > 0 \forall 0 < \varepsilon < \delta$  there exists a discrete random vector  $\xi(\varepsilon)$  over  $N$  such that  $|\Delta h_{\xi(\varepsilon)}(ij|K)| < \varepsilon$  for  $ij|K \in \mathcal{M}$  and  $|\Delta h_{\xi(\varepsilon)}(ij|K)| > \delta$  for  $ij|K \in \mathcal{K}(N) \setminus \mathcal{M}$ .

**Open problem 4.** *We would like to know whether one has  $\text{pr} = \overline{\text{pr}}$  or not.*

It is easy to see that the frame  $\overline{\text{pr}}$  is closed under copying, marginalization, intersection, and lifting. Since  $\mathcal{K}(N) \in \text{pr}(M) \subseteq \overline{\text{pr}}(M)$  if  $N \subseteq M$ , it is also closed under ascetic extension. Therefore, in light of Definition 2.20, Open problem 4 is equivalent to the question whether the probabilistic CI implication on  $\mathcal{K}(N)$  coincides with the (almost entropic)  $\overline{\text{pr}}$ -closure operation (for any  $N$ ). This was basically a question raised in a private communication by our colleague Rostislav Matveev. We incline to believe that there are valid CI implications which are not true in the frame  $\overline{\text{pr}}$ .



The suspected probabilistic CI rules from [37, §V.B] which may appear to be examples of invalid implications in the  $\overline{\text{pr}}$ -frame are the rules (I:1), (I:7), (I:13), and (I:15). The reason for this guess is that to derive these implications one inevitably needs the first conditional Ingleton inequality [42] and this particular conditional inequality need not hold for almost entropic vectors [22, §V.1].

Remarks. Let us add three comments.

1. It follows from Corollary 8.3 that any CI implication valid in  $\text{st}^\diamond$ -frame can be viewed as a consequence of an unconstrained information inequality. Note, however, that *conditional information inequalities* are also studied in information theory, whose validity is restricted to and guaranteed for entropic vectors only [22]. All the so-far known probabilistic CI implications were derived from such conditional inequalities [37].
2. One of the reviewers raised the question of how many of the remaining 154 types are known to be entropic for sure. The best lower bound we are aware of is provided by *linearly representable* polymatroids which are known to be entropic. The cone generated by linear 5-polymatroids was characterized in [12] both in terms of linear inequalities and in terms of its extreme rays: they computed 7943 of such extreme rays falling into 162 permutational types. Of these 162 types exactly 138 coincide with extreme rays of the polymatroidal cone. Thus there are only 16 extreme rays of the 5-polymatroid cone for which it is unknown whether they are entropic or not. We provide the corresponding coatoms of  $\text{st}^\diamond(5)$  on our supplementary materials website.
3. Progress on the remaining 16 types can be made using stronger probabilistic rules from [37, §V.B]. This is indeed the case: we found further 11 types which cannot be probabilistic because their induced CI models  $\mathcal{M}$  do not satisfy the (extended version of the) probabilistic rule (I:2) saying that

$$ij|M, ik|\ell M, k\ell|iM, k\ell|jM \in \mathcal{M} \Rightarrow ik|M, k\ell|M \in \mathcal{M}.$$

But only 5 of these 11 types can reliably be excluded from being almost entropic as they do not satisfy the above implication with  $M = \emptyset$ ; these finer conclusions follow from the (method of) rule derivation in [37, §V.B]. We leave this observation for future work.

## 9. Conclusions

In this paper, we adapted the self-adhesivity concept, introduced originally in the context of polymatroids, to abstract CI models and explored its algebraic and order-theoretic properties. Our intention was to use existing axioms for probabilistic CI to derive stronger axioms through self-adhesion. Unlike a former paper [4], in which this strategy was applied to gaussoids, this paper emphasized the CI frames which are closed under set-theoretical intersection. Consequent lattice structure of families of models is the key to our computational advances. In particular, using state-of-the-art SAT solvers [44], we were easily able to compute  $\text{sg}^\diamond(N)$ ,  $\text{st}^\diamond(N)$ ,  $\text{gr}^\diamond(N)$  and  $\text{cgr}^\diamond(N)$  for  $|N| \leq 4$  and to derive axiomatizations for each of these families using the techniques from the formal concept analysis. On the basis of these axioms, we were able to compare the power of our approach with earlier work on CI axioms. In the context of information theory, we were able to exclude a large proportion (almost 90%) of permutational types of extreme rays of the polymatroidal cone for  $|N| = 5$  from being almost entropic. These facts can in turn motivate new unconditional information inequalities.

It was also demonstrated in [4] that computations with gaussoids over 5 variables are still possible, for the number of gaussoids  $60\,212\,776 \approx 60 \times 10^6$  [5] is still manageable. As concerns the semi-graphoids, based on a few runs of a probabilistic model counter [49] with error probability  $\delta = 0.05$  we conjecture  $|\text{sg}(5)| = 12\,144\,387\,289\,848 \approx 12 \times 10^{12}$  which rules out exhaustive computations on the set of 5-semigraphoids.<sup>7</sup> Despite that, we believe that the coatoms and irreducible elements of this lattice can be obtained by a clever arrangement of the computations and their self-adhesivity status would be valuable information in the delimitation of  $\text{pr}(5)$ . Thus, the following remains our goal.

**Open problem 5.** We would like to find all coatoms and irreducible elements of  $(\text{sg}(N), \subseteq)$  in case  $|N| = 5$ .

Note in this context that, according to a counter-example from [20], in case  $|N| = 5$ , coatoms of  $(\text{sg}(N), \subseteq)$  exist which are outside  $\text{st}(N)$ .

We also verified computationally that  $\text{sg}^\diamond(N) = \text{sg}^{\diamond\diamond}(N)$  in case  $|N| = 4$ . Computing this “second-order self-adhesion” required a significant implementation effort and took much longer than any of our other computations: while the results on single self-adhesions documented in Appendix C took roughly 18 min in total to compute, the double self-adhesion  $\text{sg}^{\diamond\diamond}(4)$  required 6 h. To explain this sudden increase, recall from Appendix B that computing the closure operator with respect to  $\text{f}^\diamond(N)$  requires the closure for  $\text{f}^\diamond(M)$  for  $|N| \leq |M| \leq 2|N| - 2$ , which in turn is computed using the closure operators  $\text{f}(O)$  for  $|M| \leq |O| \leq 2|M| - 2$ . In the case of semi-graphoids over  $|N| = 4$ , this requires semi-graphoid closures over ground sets of size up to 10. At the maximum  $|O| = 10$ , the corresponding Boolean formula used to compute closures has 184 320 clauses in 11 520 variables. We expect that  $\text{sg}^\diamond(N) = \text{sg}^{\diamond\diamond}(N)$  does not hold for  $|N| = 5$ , but an exhaustive computation is out of our reach.

<sup>7</sup> The exact number of 5-semigraphoids remains unknown because the exact model counters quickly run out of memory on the axioms of 5-semigraphoids.



In this context, it may be worthwhile to study in future some more advanced notions of self-adhesivity, which allow for a different growth in the number of ground set elements over which the closures have to be computed than iterated self-adhesion (which doubles the number of elements each time). Such notions were proposed, again in the context of polymatroids, by Csirmaz [8], who called the linear information inequalities they implied *book inequalities*. Formally, one can fix any  $k \geq 1$  and take a CI model  $\mathcal{M} \in \mathfrak{f}(N)$  and  $L \subseteq N$ . One asks whether there exists a sequence  $\mathcal{M}_0 := \mathcal{M}, \mathcal{M}_1, \dots, \mathcal{M}_k$  where  $\mathcal{M}_i$  on ground set  $N_i$  is an  $L$ -copy of  $\mathcal{M}_0$  and  $N_i \cap N_j = L$  for all pairs  $i \neq j$  and such that there exists  $\overline{\mathcal{M}} \in \mathfrak{f}\left(\bigcup_{i=0}^k N_i\right)$  which restricts to  $\mathcal{M}_i$  on  $N_i$  and satisfies  $N_{i_1} \dots N_{i_p} \perp\!\!\!\perp N_{j_1} \dots N_{j_q} \mid L \ [\overline{\mathcal{M}}]$  for any pair of disjoint subsets  $\{i_1, \dots, i_p\}$  and  $\{j_1, \dots, j_q\}$  of  $\{0, \dots, k\}$ . For  $k = 1$ , one recovers our notion of self-adhesivity. The computational advantages of CI frames whose CI families are lattices can also be exploited in this setting, and the analogue to Matúš’s fundamental result [29] on the closedness of  $\text{pr}$  under such generalized self-adhesivity follows from the same ideas as presented in our proof of Lemma 4.7.

We end with an open question about finite axiomatizability. We have seen in Section 7.1 that the semi-graphoidal frame  $\text{sg}$  is special in the sense that the canonical implicational basis for  $(\text{sg}(N), \subseteq)$  for any  $|N|$  follows one scheme only. One can interpret this by saying that semi-graphoids have a finite “axiomatization” (with a single axiom).

**Open problem 6.** *We would like to know whether  $\text{sg}^\diamond$  admits a finite “axiomatic” characterization, too.*

### Acknowledgements

Most of this work was inspired by the research of our late colleague Fero Matúš. Tobias Boege was partially supported by the Academy of Finland, Finland grant number 323416 and by the Wallenberg Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, Sweden. The authors discussed in person the topic of this paper at two meetings in 2022, one of them was the DAGSTUHL seminar 22301 “Algorithmic Aspects of Information Theory”. We also thank both reviewers of DAM journal for the valuable comments which helped to improve the quality of presentation of our results.

### Appendix A. Algorithm to compute the canonical basis

This is a pseudo-code of the algorithm we used to obtain the canonical generator  $\mathbb{G}_{\mathcal{F}}$  on basis of a Moore family  $\mathcal{F}$  of subsets of a finite set  $X$  or its closure operation  $\text{cl}_{\mathcal{F}}$ :

1. choose the maximal parameter  $m$ ,  $-1 \leq m \leq |X|$ , such that any set  $Y \subseteq X$  with  $|Y| \leq m$  is closed (that is,  $Y \in \mathcal{F}$ ); put  $\mathbb{G} := \emptyset \subseteq \mathcal{P}(X) \times \mathcal{P}(X)$ ,
2. unless  $m = |X|$  raise  $m := m + 1$ ; if  $m = |X|$  then put  $\mathbb{G}_{\mathcal{F}} := \mathbb{G}$  and exit,
3. test  $Y \subseteq X$  with  $|Y| = m$  whether they are pseudo-closed (Definition 2.14)<sup>8</sup>; enlarge  $\mathbb{G} := \mathbb{G} \cup \{(Y \rightarrow \text{cl}_{\mathcal{F}}(Y) \setminus Y) : Y \in \wp(\mathcal{F}), |Y| = m\}$ ,
4. then compute  $\mathbb{G}^\triangleright$  and branch:
  - (a) if  $|\mathbb{G}^\triangleright| > |\mathcal{F}|$  then go to 2,
  - (b) if  $|\mathbb{G}^\triangleright| = |\mathcal{F}|$  then  $\mathbb{G}_{\mathcal{F}} := \mathbb{G}$  and exit.

Note that in Step 1, when  $\mathcal{F} = \mathfrak{f}(N)$  for a suitable CI family  $\mathfrak{f}$ , one typically has  $m = 1$ . By the construction of  $\mathbb{G}$ , we know that  $\mathbb{G} \subseteq \mathcal{F}^\triangleleft$  and  $\mathcal{F} \subseteq \mathbb{G}^\triangleright$  in every phase of the algorithm. In Step 4(b), we are sure that  $\mathbb{G}$  is the implicational generator of  $\mathcal{F}$  by the criterion mentioned above Example 2.13.

What is described above is just one of basic versions of the procedure to compute the canonical basis. There are surely various more refined and efficient versions of this procedure [15, §3.3.2]; the paper [11] was particularly devoted to the comparison of computational complexities of such algorithms.

### Appendix B. Algorithm to test iterated self-adhesivity

Consider a CI frame  $\mathfrak{f}$  closed under copying, marginalization, intersection, lifting, tight replication, and ascetic extension. The basis of the algorithm is a subroutine to compute the  $\mathfrak{f}^\diamond$ -closure of a given model  $\hat{\mathcal{R}} \in \mathfrak{f}(M)$ , which is a modification of the procedure in Lemma 5.4.

Subroutine  $\mathfrak{f}^\diamond(\hat{\mathcal{R}})$  [the input is  $\hat{\mathcal{R}} \in \mathfrak{f}(M)$ ]

1. put  $\mathcal{R} := \hat{\mathcal{R}}$
2. apply to  $\mathcal{R}$  the following cycle governed by  $L \subseteq M$ ,  $|L| = 2, \dots, |M| - 2$ :
  - (a) compute the self-adhesive closure  $\mathfrak{f}^\diamond(\mathcal{R}|L)$  as described in Lemma 4.5,
  - (b) test whether  $\mathfrak{f}^\diamond(\mathcal{R}|L) \setminus \mathcal{R} \neq \emptyset$ :

<sup>8</sup> Note that pseudo-closed strict subsets of  $Y$  are encoded in current version of  $\mathbb{G}$ .

- if yes then put  $\mathcal{R} := f^\circ(\mathcal{R}|L)$ , exit the cycle, and go to 2,
- if not then continue within the cycle with the next set  $L$ ,

3. put  $f^\circ(\hat{\mathcal{R}}) := \mathcal{R}$  and exit.

Note that in Step 2, by [Corollary 4.12](#) and [Lemma 4.16](#), we need not test sets  $L$  of cardinalities 0, 1,  $|M| - 1$ , and  $|M|$ . As concerns updating of  $\mathcal{R}$  in Step 2(b), by [Definition 4.3](#),  $f^\circ(\mathcal{R}|L) \in f(M)$ .

The main procedure is analogous and differs in small details only. The main difference is that [Lemma 4.5](#) is this time applied to  $f^\circ$  instead of  $f$ . Note that this modification is possible thanks to [Lemma 5.1\(iii\)](#).

Main procedure  $f^{\circ\circ}(\hat{\mathcal{M}})$  [the input is  $\hat{\mathcal{M}} \in f^\circ(N)$ , where  $|N| = 4$ ]

1. put  $\mathcal{M} := \hat{\mathcal{M}}$
2. apply to  $\mathcal{M}$  the following cycle governed by  $E \subseteq N$ ,  $|E| = 2, 3$ :
  - (a) compute the self-adhesive closure  $f^{\circ\circ}(\mathcal{M}|E)$  using [Lemma 4.5](#), where the subroutine is called to compute the  $f^\circ$ -closures,
  - (b) test whether  $f^{\circ\circ}(\mathcal{M}|E) \setminus \mathcal{M} \neq \emptyset$ :
    - if yes then put  $\mathcal{M} := f^{\circ\circ}(\mathcal{M}|E)$ , exit the cycle, and go to 2,
    - if not then continue within the cycle with the next set  $E$ ,
3. if  $\mathcal{M} \neq \hat{\mathcal{M}}$  then  $\hat{\mathcal{M}} \notin f^{\circ\circ}(N)$  and its closure is  $f^{\circ\circ}(\hat{\mathcal{M}}) := \mathcal{M}$ ; exit.

Note that, in Step 2, by [Lemma 5.1\(i\)\(ii\)](#) and [Corollary 4.12](#), we need not test sets  $E$  of cardinalities 0,1, and 4. On the other hand, since we are not sure whether  $f^\circ$  is closed under tight replication (see [Open problem 1](#)) we have to test sets  $E$  of cardinality 3. So, the number of sets  $E$  in line 2 is at most 10.

To check double self-adhesivity over 4 variables one needs to apply the subroutine for  $|M| = 5, 6$ , in which case the number of sets  $L$  in step 2 of the subroutine is at most 50. Also, to compute  $f^\circ(\mathcal{R}|L)$  one applies the  $f$ -closures over at most 10 variables.

**Appendix C. Table of our results in case of 4 variables**

Family	Models	in types	Irreducibles	in types	Coatoms	in types
$sg(N)$	26 424	1512	181	20	37	10
$sg^\circ(N)$	23 190	1352	385	29	31	9
$st(N)$	22 108	1285	37	10	37	10
$st^\circ(N)$	20 968	1224	85	13	31	9
$gr(N)$	6482	421	408	32	14	4
$gr^\circ(N)$	6482	421	408	32	14	4
$gr^*(N)$	6482	421	408	32	14	4
$(gr^*)^\circ(N)$	6182	404	428	33	14	4
$cgr(N)$	2084	157	470	30	6	1
$cgr^\circ(N)$	2084	157	470	30	6	1

**Appendix D. Irreducible self-adhesive structural models**

Here we give the list of permutational types of (meet) irreducible elements of the lattice  $(st^\circ(N), \subseteq)$  in case  $|N| = 4^9$ :

- I. {  $[ijk, \ell|\emptyset], [ij, k|\emptyset], [ij, k|\ell]$  },
- II. {  $[ijk, \ell|\emptyset], ij|k, ij|k\ell, ik|j, ik|j\ell, jk|i, jk|i\ell$  },
- III. {  $[ij, k|\ell], [ij, \ell|k], ij|k, ij|\ell, ij|k\ell, ik|j, i\ell|j, jk|i, j\ell|i, k\ell|i, k\ell|j, k\ell|ij$  },
- IV. {  $[ijk, \ell|\emptyset], ij|\emptyset, ij|\ell, ik|\emptyset, ik|\ell, jk|\emptyset, jk|i\ell$  },
- V. {  $[ij, k|\emptyset], [ij, \ell|\emptyset], ij|\emptyset, ij|k, ij|\ell, ik|\ell, i\ell|k, jk|\ell, j\ell|k, k\ell|\emptyset, k\ell|i, k\ell|j$  },
- VI. {  $[ij, k|\ell], [ij, \ell|k], ij|\emptyset, ik|\emptyset, i\ell|\emptyset, jk|\emptyset, j\ell|\emptyset, k\ell|ij$  },
- VII. {  $ij|\emptyset, ij|k\ell, ik|\emptyset, ik|j\ell, i\ell|\emptyset, i\ell|jk, jk|\emptyset, jk|i\ell, j\ell|\emptyset, j\ell|ik, k\ell|\emptyset, k\ell|ij$  },
- VIII. {  $[ij, k|\ell], ij|\emptyset, ij|\ell, ij|k\ell, ik|\emptyset, i\ell|jk, jk|\emptyset, j\ell|ik, k\ell|ij$  },

<sup>9</sup> Recall that we apply the notational convention from [Definition 2.3](#).

IX.  $\{ ij|k, ij|\emptyset, ij|k, ij|kl, ik|j\ell, il|\emptyset, jk|i\ell, j\ell|\emptyset, kl|\emptyset \}$ ,

X.  $\{ ij|\emptyset, ij|k, ij|\ell, ik|\ell, jk|\ell, kl|i, kl|j, kl|ij \}$ ,

XI.  $\{ ij|\emptyset, ij|k, ij|\ell, ik|\ell, j\ell|k, kl|i, kl|j, kl|ij \}$ ,

XII.  $\{ ij|\emptyset, ij|k, ij|\ell, ik|\ell, il|k, jk|\ell, kl|j, kl|ij \}$ ,

XIII.  $\{ ij|\emptyset, ij|k, ij|\ell, ik|\ell, il|k, jk|\ell, j\ell|k, kl|ij \}$ .

The types I. – IX. involve all 31 coatoms of the lattice. Their notation follows the classification used in [37, Appendix B]. The actual numbers of involved elementary CI statements are as follows:

- I. consists of 20 CI statements,
- II. – V. consists of 18 CI statements,
- VI. consists of 14 CI statements,
- VII. – IX. consists of 12 CI statements, and
- X. – XIII. consists of 8 CI statements.

### Data availability

The data is available on the internet. Link is featured in the paper.

### References

- [1] Sanjeev Arora, Boaz Barak, Computational Complexity. A Modern Approach, Cambridge University Press, 2009, <http://dx.doi.org/10.1017/CBO9780511804090>.
- [2] Garrett Birkhoff, Lattice Theory, third ed., in: Colloq. Publ., Am. Math. Soc., vol. 25, American Mathematical Society, Providence, 1995.
- [3] Tobias Boege, The Gaussian Conditional Independence Inference Problem (Ph.D. thesis), OVGU Magdeburg, 2022, <http://dx.doi.org/10.25673/86275>.
- [4] Tobias Boege, Selfadhesivity in Gaussian conditional independence structures, Int. J. Approx. Reason. 163 (2023) 11, <http://dx.doi.org/10.1016/j.ijar.2023.109027>, Id/No 109027.
- [5] Tobias Boege, Alessio D’Alì, Thomas Kahle, Bernd Sturmfels, The geometry of gaussoids, Found. Comput. Math. 19 (4) (2019) 775–812, <http://dx.doi.org/10.1007/s10208-018-9396-x>.
- [6] Janneke H. Bolt, Two generalizations of the semi-graphoid rule of probabilistic independence and more, Int. J. Approx. Reason. 161 (2023) 16, <http://dx.doi.org/10.1016/j.ijar.2023.108985>, Id/No 108985.
- [7] Janneke H. Bolt, Linda C. van der Gaag, Generalized rules of probabilistic independence, in: Symbolic and Quantitative Approaches to Reasoning with Uncertainty. 16th European Conference, ECSQARU 2021, Prague, Czech Republic, September 21–24, 2021. Proceedings, Springer, Cham, 2021, pp. 590–602, [http://dx.doi.org/10.1007/978-3-030-86772-0\\_42](http://dx.doi.org/10.1007/978-3-030-86772-0_42).
- [8] László Csirmaz, Book inequalities, IEEE Trans. Inform. Theory 60 (11) (2014) 6811–6818, <http://dx.doi.org/10.1109/TIT.2014.2352273>.
- [9] László Csirmaz, One-adhesive polymatroids, Kybernetika 56 (5) (2020) 886–902, <http://dx.doi.org/10.14736/kyb-2020-5-0886>.
- [10] A. Philip Dawid, Conditional independence in statistical theory, J. R. Stat. Soc. Ser. B Stat. Methodol. 41 (1979) 1–31.
- [11] Felix Distel, Bariš Sertkaya, On the complexity of enumerating pseudo-intents, Discrete Appl. Math. 159 (6) (2011) 450–466, <http://dx.doi.org/10.1016/j.dam.2010.12.004>.
- [12] Randall Dougherty, Chris Freiling, Kenneth Zeger, Linear rank inequalities on five or more variables, 2010, [arXiv:0910.0284v3](https://arxiv.org/abs/0910.0284v3).
- [13] Randall Dougherty, Chris Freiling, Kenneth Zeger, Non-Shannon information inequalities in four random variables, 2011, [arXiv:1104.3602v1](https://arxiv.org/abs/1104.3602v1).
- [14] Satoru Fujishige, Submodular functions and optimization, second ed., in: Ann. Discrete Math., vol. 58, Elsevier, Amsterdam, 2005.
- [15] Bernhard Ganter, Sergei Obiedkov, Conceptual Exploration, Springer, Berlin, 2016, <http://dx.doi.org/10.1007/978-3-662-49291-8>.
- [16] Bernhard Ganter, Rudolf Wille, Formal Concept Analysis – Mathematical Foundations, Springer, Berlin, 1999.
- [17] Dan Geiger, Judea Pearl, Logical and algorithmic properties of conditional independence and graphical models, Ann. Statist. 21 (4) (1993) 2001–2021, <http://dx.doi.org/10.1214/aos/1176349407>.
- [18] Jean-Louis Guigues, Vincent Duquenne, Familles minimales d’implications in-formatives résultant d’un tableau de données binaires, Math. Sci. Hum. 95 (1986) 5–18.
- [19] Laigang Guo, Raymond W. Yeung, Xiao-Shan Gao, Proving information inequalities and identities with symbolic computation, IEEE Trans. Inform. Theory 69 (8) (2023) 4799–4811, <http://dx.doi.org/10.1109/TIT.2023.3263178>.
- [20] Raymond Hemmecke, Jason Morton, Anne Shiu, Bernd Sturmfels, Oliver Wienand, Three counter-examples on semi-graphoids, Combin. Probab. Comput. 17 (2) (2008) 239–257, <http://dx.doi.org/10.1017/S0963548307008838>.
- [21] Alfred Horn, On sentences which are true of direct unions of algebras, J. Symbolic Logic 16 (1951) 14–21, <http://dx.doi.org/10.2307/2268661>.
- [22] Tarik Kaced, Andrei Romashchenko, Conditional information inequalities for entropic and almost entropic points, IEEE Trans. Inform. Theory 59 (11) (2013) 7149–7167, <http://dx.doi.org/10.1109/TIT.2013.2274614>.
- [23] Radim Lněnička, František Matúš, On Gaussian conditional independence structures, Kybernetika 43 (3) (2007) 327–342, URL [www.kybernetika.cz/content/2007/3/327](http://www.kybernetika.cz/content/2007/3/327).
- [24] František Matúš, Abstract functional dependency structures, Theoret. Comput. Sci. 81 (1) (1991) 117–126, [http://dx.doi.org/10.1016/0304-3975\(91\)90319-W](http://dx.doi.org/10.1016/0304-3975(91)90319-W).
- [25] František Matúš, Ascending and descending conditional independence relations, in: Information Theory, Statistical Decision Functions, and Random Processes. Transactions of the 11th Prague Conference, Held from August 27–31, 1990 in Prague, Czechoslovakia, vol. B, Kluwer Academic Publishers, Dordrecht, 1992, pp. 189–200.

- [26] František Matúš, Conditional independences among four random variables II, *Combin. Probab. Comput.* 4 (4) (1995) 407–417, <http://dx.doi.org/10.1017/S0963548300001747>.
- [27] František Matúš, Lengths of semigraphoid inferences, *Ann. Math. Artif. Intell.* 35 (1–4) (2002) 287–294, <http://dx.doi.org/10.1023/A:1014525817725>.
- [28] František Matúš, Towards classification of semigraphoids, *Discrete Math.* 277 (1–3) (2004) 115–145, [http://dx.doi.org/10.1016/S0012-365X\(03\)00155-9](http://dx.doi.org/10.1016/S0012-365X(03)00155-9).
- [29] František Matúš, Adhesivity of polymatroids, *Discrete Math.* 307 (21) (2007) 2464–2477, <http://dx.doi.org/10.1016/j.disc.2006.11.013>.
- [30] František Matúš, Milan Studený, Conditional independences among four random variables I, *Combin. Probab. Comput.* 4 (3) (1995) 269–278, <http://dx.doi.org/10.1017/S0963548300001644>.
- [31] James G. Oxley, *Matroid theory, second edition*, *Oxf. Grad. Texts Math.*, vol. 21, Oxford University Press, Oxford, 2011.
- [32] Judea Pearl, *Probabilistic Reasoning in Intelligent Systems – Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Mateo, 1989.
- [33] Kayvan Sadeghi, Steffen Lauritzen, Markov properties for mixed graphs, *Bernoulli* 20 (2) (2014) 676–696, <http://dx.doi.org/10.3150/12-BEJ502>.
- [34] Milan Studený, Structural semigraphoids, *Int. J. Gen. Syst.* 22 (2) (1994) 207–217, <http://dx.doi.org/10.1080/03081079308935207>.
- [35] Milan Studený, *Probabilistic conditional independence structures*, in: *Inf. Sci. Stat.*, Springer, London, 2005.
- [36] Milan Studený, Basic Facts Concerning Extreme Supermodular Functions, Tech. Rep. 2359, Institute of Information Theory and Automation, Prague, 2016, Also available on <https://arxiv.org/abs/1612.06599>.
- [37] Milan Studený, Conditional independence structures over four discrete random variables revisited: conditional Ingleton inequalities, *IEEE Trans. Inform. Theory* 67 (11) (2021) 7030–7049, <http://dx.doi.org/10.1109/TIT.2021.3104250>.
- [38] Milan Studený, Remco R. Bouckaert, Tomáš Kočka, *Extreme Supermodular Set Functions Over Five Variables*, Tech. Rep. 1977, Institute of Information Theory and Automation, Prague, 2000.
- [39] Milan Studený, Tomáš Kroupa, Core-based criterion for extreme supermodular functions, *Discrete Appl. Math.* 206 (2016) 122–151, <http://dx.doi.org/10.1016/j.dam.2016.01.019>.
- [40] Simon Vilmin, *Algorithms on Closure Systems and their Representations (Ph.D. thesis)*, Université Clermont Auvergne, 2021.
- [41] Raymond W. Yeung, *Information Theory and Network Coding*, Springer, New York, 2008.
- [42] Zhen Zhang, Raymond W. Yeung, A non-Shannon-type conditional inequality of information quantities, *IEEE Trans. Inform. Theory* 43 (6) (1997) 1982–1986, <http://dx.doi.org/10.1109/18.641561>.
- [43] Zhen Zhang, Raymond W. Yeung, On characterization of entropy function via information inequalities, *IEEE Trans. Inform. Theory* 44 (4) (1998) 1440–1452, <http://dx.doi.org/10.1109/18.681320>.

## Mathematical software

- [44] Armin Biere, CaDiCaL at the SAT Race 2019, in: Marijn Heule, Matti Järvisalo, Martin Suda (Eds.), *Proc. of SAT Race 2019 – Solver and Benchmark Descriptions*, in: *Department of Computer Science Series of Publications B*, vol. B-2019-1, University of Helsinki, 2019, pp. 8–9.
- [45] Tobias Boege, Clnet::Tools – A software for computations with conditional independence structures, 2024, Available at <https://github.com/Clnet>.
- [46] Winfried Bruns, Bogdan Ichim, Christof Söger, Ulrich von der Ohe, Normaliz. Algorithms for rational cones and affine monoids, 2023, Available at <https://www.normaliz.uni-osnabrueck.de>.
- [47] Ambros Gleixner, Michael Bastubbe, Leon Eifler, Tristan Gally, Gerald Gamrath, Robert Lion Gottwald, Gregor Hendel, Christopher Hojny, Thorsten Koch, Marco Lübbecke, Stephen J. Maher, Matthias Miltenberger, Benjamin Müller, Marc Pfetsch, Christian Puchert, Daniel Rehfeldt, Franziska Schlösser, Christoph Schubert, Felipe Serrano, Yuji Shinano, Jan Merlin Viernickel, Matthias Walter, Fabian Wegscheider, Jonas T. Witt, Jakob Witzig, *The SCIP Optimization Suite 6.0*, Tech. Rep. 18–26, ZIB, Takustr. 7, 14195 Berlin, 2018.
- [48] Ambros Gleixner, Daniel Steffy, Kati Wolter, *Iterative Refinement for Linear Programming*, Tech. Rep. 15–15, ZIB, Takustr. 7, 14195 Berlin, 2015.
- [49] Shubham Sharma, Subhajit Roy, Mate Soos, Kuldeep S. Meel, GANAK: A scalable probabilistic exact model counter, in: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization*, 2019, pp. 1169–1176, <http://dx.doi.org/10.24963/ijcai.2019/163>.
- [50] Takahisa Toda, Takehide Soh, Implementing efficient all solutions SAT solvers, *J. Exp. Algorithmics* 21 (2016) 1.12:1–1.12:44, <http://dx.doi.org/10.1145/2975585>.
- [51] Wolfram Research, Inc., *Mathematica*, 2023, Champaign, IL. Version 13.3.