ON CHAIN GRAPH MODELS FOR DESCRIPTION OF CONDITIONAL INDEPENDENCE STRUCTURES[1]

By Milan Studený and Remco R. Bouckaert

Academy of Sciences of Czech Republic and Utrecht University

A *chain graph* (CG) is a graph admitting both directed and undirected edges with (partially) directed cycles forbidden. It generalizes both the concept of undirected graph (UG) and the concept of directed acyclic graph (DAG). A chain graph can be used to describe efficiently the conditional independence structure of a multidimensional discrete probability distribution in the form of a *graphoid*, that is, in the form of a list of statements "X is independent of Y given Z" obeying a set of five properties (axioms).

An input list of independency statements for every CG is defined and it is shown that the classic *moralization criterion* for CGs embraces exactly the graphoid closure of the input list. A new direct *separation criterion* for reading independency statements from a CG is introduced and shown to be equivalent to the moralization criterion. Using this new criterion it is proved that for every CG there exists a strictly positive discrete probability distribution that embodies exactly the independency statements displayed by the graph. Thus, both criteria are shown to be complete and the use of CGs as tools for description of conditional independence structures is justified.

1. Introduction. Using graphs to describe conditional independence structures among random variables has a long and rich tradition. One can distinguish two classic approaches (for details see the book [22]): using *undirected graphs* (UGs), also called Markov networks, or using *directed acyclic graphs* (DAGs), also named Bayesian networks or (probabilistic) influence diagrams. The aim was to describe efficiently independency models in the form of lists of statements "X is independent of Y given Z", where X, Y, Z are disjoint sets of variables. Such structures can arise in several calculi for dealing with uncertainty in artificial intelligence: in probabilistic reasoning, in the theory of natural conditional functions known also as kappa-calculus, in possibility theory or Dempster-Shafer's theory of evidence (for an overview see [25]) but also in the theory of relational databases. Of course, different calculi produce different independency models, but in case of nonextreme knowledge representation they share five basic prop-

[1]This work was supported by the grant of the AVČR n. 275105, by the grants GAČR n. 201/94/0471 and n. 201/98/0478, and by the grant MŠMT n. VŠ96008.

AMS 1991 *subject classifications.* Primary 62H99; secondary 62H05, 68R10.

Key words and phrases. Chain graph, conditional independence, Markovian distribution, input list, c-separation, completeness.

erties which define the class of *graphoids* [22]. Note that the importance of those formal (graphoid) properties in statistics was emphasized by Dawid [8].

Graphoids can sometimes be described graphically. Thus, every UG defines by means of a separation criterion an independency model which is a graphoid. The use of UGs in probabilistic reasoning was justified by the result from [12], where every such UG-model was shown to be a probabilistic independency model. Nevertheless, many graphoids (even probabilistic models) have no UG representation (that is, are not UG-models). Therefore Pearl [22] proposed to approximate graphoids by their contained UG-models (I-maps) and showed that for every graphoid M there exists a unique maximal UG-model contained in M, called the *minimal I-map of* M.

The evolution of DAG-models was more complicated. Originally, DAGs were used to describe recursive factorizations of probability distributions. But such a factorization is equivalent to the requirement that the considered distribution satisfies a set of independency statements known as the *causal input list*. Nevertheless, the distribution usually satisfies many other independency statements outside the causal input list. Much effort was exerted to find a graphical criterion that makes it possible to read from a DAG all independency statements necessarily valid in the factorizable distribution. In fact, two equivalent criteria were found. Lauritzen *et. al.* [18] generalized an incomplete criterion from [13] and formulated a *moralization criterion* where testing consists of 3 steps: restriction of the DAG to a certain set of nodes, transforming it properly to an UG (called the *moral graph*), and using the separation criterion for UGs with respect to the moral graph. The group around Pearl developed a direct *separation criterion* [10], based on the concept of *d-separation* (directional separation) for paths in DAGs. It was shown that the criteria are equivalent [18] and that they give exactly the graphoid closure of the causal input list [28]. Finally, the criteria were shown to be complete for probabilistic reasoning by showing that every independency model defined by the separation criterion is a probabilistic independency model [11]. Thus, DAG-models were established and their use in probabilistic reasoning was justified. As in the case of UGs Pearl [22] considered the problem of inner approximation of graphoids by DAG-models. In contrast to the case of UGs several maximal DAG-models contained in a graphoid may exist. In fact, any ordering of variables can generate a minimal I-map of a graphoid; the corresponding construction is given in [28].

This paper deals with *chain graphs* (CGs) which allow both directed and undirected edges. This class of graphs, introduced by Lauritzen and Wermuth [15], generalizes both UGs and DAGs. To establish the semantics of CGs one should associate an independency model to every CG. Some steps were already made. Lauritzen and Wermuth [16] intended to use CGs to describe independency models for strictly positive probability distributions and introduced the concept of the *chain Markov property* which is analogous to the concept of causal input list for DAGs. Lauritzen and Frydenberg [17, 9] generalized the concept of moral graph and introduced a *moralization criterion* for reading independency statements from a CG. Frydenberg [9] characterized CGs with the same Markov property (that is producing the same CG-model) and Andersson, Madigan and Perlman [3] used special CGs to represent uniquely classes of Markov equivalent DAGs. Whittaker [31] in his book gave several examples of the use of CGs, and other recent works also deal with them [6, 20, 23, 30], the most comprehensive account is provided by the book [19]. Several results proved here were already presented (without proof) in our previous conference contribution [5].

An alternative approach to the generalization of UGs and DAGs was started by Cox and Wermuth [7] who introduced a wider class of *joint-response chain graphs* which allow also 'dashed' directed and undirected edges in addition to the classic 'solid' directed and undirected edges treated in this paper. Andersson, Madigan and Perlman [1] introduced an *alternative Markov property* to give an interpretation to those joint-response CGs which combine dashed directed edges with solid undirected edges (of course, another independency model is associated with those CGs). The different interpretations were compared by Richardson [23].

In this paper some of the above mentioned results concerning UGs and DAGs are extended to the case of CGs. In the next section, we introduce terminology and recall some known basic properties. In Section 3 we define the concept of input list of independency statements for CGs and show that the independency statements which can be read from the graph by means of the moralization criterion are exactly those derivable from the input list by means of the graphoid properties. In Section 4 we introduce a new direct graphical criterion for reading independency statements from a CG, based on the concept of *c-separation* (chain separation) for routes in a CG, and show its equivalence to the moralization criterion. In Section 5 we prove the completeness of the separation criterion, that is whenever a statement (that is a potential independency statement) cannot be read from the graph according to that criterion, then there exists a strictly positive Markovian distribution (that is satisfying the input list) which does not satisfy the considered statement. The Appendix (Section 6) contains technical proofs of auxiliary results, important for the completeness proof. In the Conclusions (Section 7) we summarize the results and derive the strong completeness result which says that for every CG there exists a strictly positive discrete probability distribution satisfying exactly those independency statements that can be read from the graph. We discuss the potential usefulness of CG models and formulate open questions in that section, too.

2. Preliminaries In this section, we review basic concepts and some of their properties. First, we separately give definitions of probabilistic and graphical concepts. Then we recall the definitions concerning the graphical representation of independency models.

CONVENTION 1. Throughout the paper, we sometimes use a shortened notation: juxtaposition XY instead of $X \cup Y$ and u instead of $\{u\}$. Thus, for example $X \setminus Yu$ denotes $X \setminus (Y \cup \{u\})$.

2.1. *Independency models* Let N be a nonempty finite set of variables. Then $T(N)$ denotes the collection of all triplets $\langle X,Y|Z\rangle$ of disjoint subsets of N whose first two components X and Y are nonempty. An *independency model over* N is a subset of $T(N)$. {Note that Pearl [22] used the term 'dependency model' just for brevity. However, as we interpret triplets in $T(N)$ as independency statements, we decided to use the term 'independency model'.} The class of *elementary triplets* $E(N)$ consists of $\langle a,b|Z\rangle \in T(N)$ where $a,b \in N$ are distinct and $Z \subset N \setminus ab$.

A *potential over* a set $Y \subset N$ is specified by a collection of nonempty finite sets $\{\mathbf{X}_i; i \in Y\}$ and by a nonnegative real function R on the cartesian product $\prod_{i\in Y} \mathbf{X}_i$. We accept a natural convention that the empty cartesian product $\prod_{i\in\varnothing} \mathbf{X}_i$ is regarded as a singleton. Thus, a potential over $\varnothing$ is nothing but a nonnegative number. If $R(\mathbf{y}) > 0$ for all $\mathbf{y} \in \prod_{i\in Y} \mathbf{X}_i$, then R is called *strictly positive*. The *marginal* of R *for* a set $X \subset Y$ is a potential over X, denoted by R^X and defined by the formula $R^X(\mathbf{x}) = \sum\{R(\mathbf{x},\mathbf{w});\ \mathbf{w} \in \prod_{i\in Y\setminus X} \mathbf{X}_i\}$, where $\mathbf{x} \in \prod_{i\in X} \mathbf{X}_i$.

A potential P over N with $P^{\varnothing} = 1$, that is $\sum\{P(\mathbf{x});\ \mathbf{x} \in \prod_{i\in N} \mathbf{X}_i\} = 1$, is called a *probability distribution over* N. {Note that throughout the paper we limit ourselves to discrete probability distributions.} Evidently, the marginal of a probability distribution P over N for $X \subset N$ is a probability distribution over X, called the *marginal distribution* of P *for* X.

Let P be a probability distribution over N and $\langle X,Y|Z\rangle \in T(N)$. Then, we say that X is *conditionally independent of* Y *given* Z *with respect to* P or that P *satisfies* $\langle X,Y|Z\rangle$ if $\quad \forall\, \mathbf{x} \in \prod_{i\in X} \mathbf{X}_i \quad \mathbf{y} \in \prod_{i\in Y} \mathbf{X}_i \quad \mathbf{z} \in \prod_{i\in Z} \mathbf{X}_i$

$$P^{XYZ}(\mathbf{x},\mathbf{y},\mathbf{z}) \cdot P^{Z}(\mathbf{z}) = P^{XZ}(\mathbf{x},\mathbf{z}) \cdot P^{YZ}(\mathbf{y},\mathbf{z}),$$

and write $X \perp\!\!\!\perp Y \mid Z\ [P]$. {Here we follow Dawid's original notation [8].} The statement $X \perp\!\!\!\perp Y \mid Z\ [P]$ is called an *independency statement.*

Each probability distribution P over N defines an independency model over N, namely the model $\{\, \langle X,Y|Z\rangle \in T(N);\ X \perp\!\!\!\perp Y \mid Z\ [P]\,\}$ called the *independency model induced by* P. An independency model over N is then called *probabilistic* if it is induced by some probability distribution over N.

Let $\mathcal{P}$ be a class of probability distributions over N, for example the class of class of strictly positive probability distributions over N. Then having $L \subset T(N)$ and $u \in T(N)$ we will say that L *probabilistically implies* u *w.r.t.* $\mathcal{P}$ if every independency model induced by a $P \in \mathcal{P}$ containing L also contains u. Having a set $L \subset T(N)$ its *probabilistic closure w.r.t.* $\mathcal{P}$ is the set of all triplets in $T(N)$ which are probabilistically implied by L w.r.t. $\mathcal{P}$.

An independency model M over N is called a *semigraphoid* if it satisfies the following properties:

$\langle X,Y|Z\rangle \longrightarrow \langle Y,X|Z\rangle$	symmetry
$\langle X,YW|Z\rangle \longrightarrow \langle X,W|Z\rangle$	decomposition
$\langle X,YW|Z\rangle \longrightarrow \langle X,Y|WZ\rangle$	weak union
$[\langle X,Y|WZ\rangle \ \&\ \langle X,W|Z\rangle] \longrightarrow \langle X,YW|Z\rangle$	contraction

where $W, X, Y, Z \subset N$ are pairwise disjoint. One should understand the properties as follows: if M contains the triplets before the long arrow, then M contains also the triplet after the long arrow. A semigraphoid M is called a *graphoid* if it moreover satisfies:

$[\langle X, Y | WZ\rangle \ \& \ \langle X, W | YZ\rangle] \ \longrightarrow \ \langle X, YW | Z\rangle$ intersection.

Having a set $L \subset T(N)$ its *graphoid closure*, denoted by $\mathrm{gr}(L)$, consists of all triplets in $T(N)$ derivable from L by means of consecutive application of graphoid properties. Analogously, the *semigraphoid closure* is defined. It is well-known that every probabilistic independency model is a semigraphoid and every independency model induced by a strictly positive distribution is a graphoid, see [8]. Specifically, for $L \subset T(N)$ every triplet from $\mathrm{gr}(L)$ belongs to the probabilistic closure of L w.r.t. the class of strictly positive probability distributions.

2.2. *Graphical concepts* In this paper we consider graphs with (possibly) three different kinds of edges, namely with *undirected*, *directed* and *virtual* edges. Supposing N is a finite nonempty set of *nodes* a (*virtual*) *graph over* N is specified as follows. For every ordered pair $[u, v]$ of distinct nodes $u, v \in N$ exclusively one of the following five possibilities is determined:

- there is a *line* (= an undirected edge) between u and v, denoted by $u - v$,
- there is an *arrow* (= a directed edge) from u to v, denoted by $u \to v$,
- there is an *arrow* from v to u, denoted by $u \leftarrow v$,
- there is a *virt* (= a virtual edge) between u and v, denoted by $u \cdots v$,
- there is *no edge* between u and v in the graph.

Moreover, it is understood that whenever one of those possibilities is determined for a pair $[u, v]$, then it is also determined for the dual pair $[v, u]$ according to the following rules: $v - u$ iff $u - v$, $v \to u$ iff $u \leftarrow v$, $v \leftarrow u$ iff $u \to v$, $v \cdots u$ iff $u \cdots v$, and there is no edge between v and u iff there is no edge between u and v. The reader may notice that our terminology reflects the convention exactly: we say 'between' in the case of a symmetric relationship and 'from - to' in the case of an asymmetric relationship. The chosen mathematical notation also reflects the way that the edges are represented in pictures. We say that a two-element subset $\{u, v\}$ of N (i.e. an unordered pair) is an edge in the graph if either $u - v$ or $u \to v$ or $u \leftarrow v$ or $u \cdots v$. It is evident from the definition that at most one edge can occur between two distinct nodes in our graphs (for example it is impossible to have simultaneously $u \to v$ and $u - v$). Similarly, no edge returning to the same node exist in our graphs (for example, it is impossible to have $u \cdots u$). That is a principal difference from the case of more general *reciprocal graphs* [14] or from the case of general directed graphs [24]. Moreover, the reader should be aware of the fact that our virts do not correspond to the 'dashed lines' used by Cox and Wermuth [7]. We use virts as an auxiliary tool, to make some proofs more elegant and easy to undestand.

An alternative (but potentially ambiguous) way to introduce a virtual graph is to say that it is a couple $G = (N, \mathcal{E})$ where N is a nonempty finite set of nodes, $\mathcal{E}$ a set of two-element subsets of N called edges and where every edge has specified its type (by this is meant that one of those corresponding four possibilities is chosen). A *hybrid graph* is a virtual graph without virts (containing lines and arrows only). An *undirected graph* (UG) is a hybrid graph without arrows (containing lines only), a *directed graph* is a hybrid graph without lines (containing arrows only). Note that the empty graph (= without edges) is simultaneously an undirected and directed graph!

A *subgraph* of a graph $G = (N, \mathcal{E})$ is a graph $H = (V, \mathcal{F})$ where $V \subset N$, $\mathcal{F} \subset \mathcal{E}$ and the type of edges in H coincides with their type in G. The *induced subgraph* of G *for* a nonempty set $T \subset N$ is a special subgraph of G, denoted by G_T, having the form $G_T = (T, \mathcal{E}_T)$, where $\mathcal{E}_T = \{\, \{u, v\} \in \mathcal{E} \,;\, u, v \in T \,\}$. The *underlying* graph of a graph G is the graph with the same set of nodes and edges as G, but every edge in the underlying graph is a line (only types of edges may differ).

A *route* in a virtual graph G is a sequence of its nodes $\rho : v_1, \ldots, v_k$, $k \geq 1$ such that $\{v_i, v_{i+1}\} \in \mathcal{E}$ for every $i = 1, \ldots, k-1$. We will also use the symbol for route ρ to denote the set of its nodes $\{v_i \,;\, i = 1, \ldots, k\}$. Note that nodes can repeat in a route, therefore the set can have less than k (distinct) nodes. By the set of edges of ρ we will understand $\{\, \{v_i, v_{i+1}\} \,;\, i = 1, \ldots, k-1 \,\}$. We will say that ρ is a route *from* a node u *to* a node w if $v_1 = u$ and $v_k = w$. A route *between* u and w is a route from u to w or from w to u. The same phrases will be applied for sets of nodes instead of single nodes. A *subroute* of ρ is a sequence of nodes $v_j, \ldots, v_t$, $1 \leq j \leq t \leq k$. We will sometimes write explicitly $\rho : a \rightarrow b — c \dashrightarrow d \leftarrow e$ when the type of edges is known. A *path* in a graph G is a route in which no node appears more than once.

Supposing ρ and σ are two routes in a virtual graph over N we will say that ρ *meets* σ (or they meet each other) if they share a node. Having a route ρ and a set of nodes $Z \subset N$, we will also say that Z *meets* ρ, or more often that ρ is *hit* by Z, if a node of ρ belongs to Z. If ρ is not hit by Z, it will be called *free* w.r.t. Z. Note that we will very often say simply that a route or subroute is hit resp. free, as the set Z will be clear from the context.

A route or subroute is *undirected* if every its edge is a line. A set of nodes S of a virtual graph G is *connected* if for every couple $u, v \in S$ there is an undirected route between u and v in G (or equivalently an undirected path). Maximal connected sets in G are called *connectivity components* of G (the maximality means that no superset of S is connected). Clearly, the connectivity component containing a node u consists of all nodes v such that there exists an undirected route (resp. path) from u to v. A *connected graph* is a graph having one connectivity component, only.

An *undirected cycle* in a graph G is a sequence of its distinct nodes $v_1, \ldots, v_k$, $k \geq 3$, such that under convention $v_{k+1} = v_1$ one has $v_i — v_{i+1}$ for $i = 1, \ldots, k$. A connected UG is called a *tree* if it contains no undirected cycles. Later we will utilize the following equivalent characterization of trees whose verification is left to the reader.

LEMMA 2.1. *Let G be an UG over N. Then the following conditions are equivalent:*

(i) G *is a tree,*
(ii) there exists just one undirected path connecting each couple of nodes in G*,*
(iii) every undirected path $s_1, \dots, s_l$*,* $l \geq 1$ *in* G *can be prolonged into a sequence of distinct nodes* $s_1, \dots, s_k$*,* $k \geq l$ *involving all nodes in* N *such that* $\forall\, 1 < i \leq k$ *there exists exactly one* $1 \leq j < i$ *with* $s_i — s_j$ *in* G*.*

A *directed pseudocycle* in a graph G is a sequence of its nodes $v_1, \dots, v_k$, $k \geq 3$, such that under convention $v_{k+1} = v_1$ one has $v_i \to v_{i+1}$ or $v_i — v_{i+1}$ for $i = 1, \dots, k$ and surely $v_j \to v_{j+1}$ for at least one $j \in \{1, \dots, k\}$. If it consists of distinct nodes it is called a *directed cycle*.

A *directed acyclic graph* (DAG) is a directed graph without directed cycles. {In fact, from the grammatical point of view, the correct phrase is 'acyclic directed graph'. However, we prefer the widely accepted abbreviation 'DAG'.} A *chain graph* (CG) is a hybrid graph without directed cycles. A *chain* for a hybrid graph G over N is a partition of N into ordered disjoint (nonempty) subsets $B_1, \dots, B_n$, $n \geq 1$ called *blocks* such that the types of edges in G are determined as follows:

- if $\{u, v\}$ is an edge in G with $u, v \in B_i$ then $u — v$,
- if $\{u, v\}$ is an edge in G with $u \in B_i, v \in B_j, i < j$ then $u \to v$.

We leave it to the reader to verify the following lemma, which justifies the terminology 'chain graph'. One can refer to [9] for some implications or to Lemma 2.1 in [26].

LEMMA 2.2. *Let G be a hybrid graph. Then the following conditions are equivalent:*

(i) G *is a chain graph,*
(ii) G *has no directed pseudocycles,*
(iii) there exists a chain for G*,*
(iv) the set of connectivity components of G *can be ordered such that it forms a chain.*

Let $G = (N, \mathcal{E})$ be a virtual graph and B a set of nodes in G. Then
$\mathrm{pa}_G(B) = \{u \in N \setminus B\,;\, u \to v \text{ for some } v \in B\}$ is the set of *parents* of B,
$\mathrm{ch}_G(B) = \{u \in N \setminus B\,;\, u \leftarrow v \text{ for some } v \in B\}$ the set of *children* of B,
$\mathrm{ne}_G(B) = \{u \in N \setminus B\,;\, u — v \text{ for some } v \in B\}$ is the set of *neighbours* of B,
$\mathrm{ad}_G(B) = \{u \in N \setminus B\,;\, \{u, v\} \in \mathcal{E} \text{ for some } v \in B\}$ the set of *adjacents* of B, and $\mathrm{bd}_G(B) = \mathrm{pa}_G(B) \cup \mathrm{ne}_G(B)$ the *boundary* of B.
A route $\pi : v_1, \dots, v_k$, $k \geq 1$ is *descending* if $\forall i = 1, \dots, k-1$ either $v_i \to v_{i+1}$ or $v_i — v_{i+1}$. If there exists a descending path from a node u to a node v, then v is a *descendant* of u, or dually u is an *ancestor* of v. The symbol $\mathrm{ds}_G(B)$ denotes the set of descendants of nodes in a set B, $\mathrm{an}_G(B)$ denotes the set of ancestors

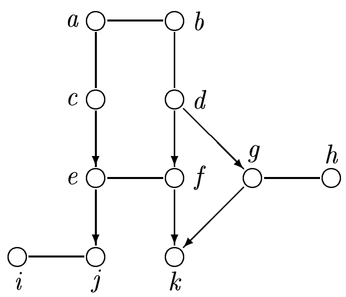

FIG. 1. *Example of a chain graph from* [9].

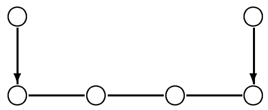

FIG. 2. *A complex of degree 4.*

of nodes in B. We will omit the symbol of the graph G if it will be clear from the context.

Note that for every connected set B in a CG (especially one node) the sets B, $\mathrm{pa}(B)$, $\mathrm{ne}(B)$, and $\mathrm{ch}(B)$ are pairwise disjoint. If nodes u and v belong to the same connectivity component of a virtual graph, then $\mathrm{ds}(u) = \mathrm{ds}(v)$ and $\mathrm{an}(u) = \mathrm{an}(v)$. Conversely, for every node u in a CG G the set $\mathrm{ds}_G(u) \cap \mathrm{an}_G(u)$ is the connectivity component of G containing u.

A set of nodes X is *ancestral* (in G) if it contains ancestors of its nodes, that is, $\mathrm{an}_G(X) \subset X$. {Note that Frydenberg [9] used the adjective 'anterior' instead of 'ancestral' and gave another equivalent definition, which is immaterial in the context of this paper.} Let us mention that an ancestral set contains with every node the connectivity component containing the node.

EXAMPLE 2.1. Figure 1 shows a chain graph G taken from [9] to illustrate some graphical concepts. In G, we have $\mathrm{pa}(e) = \{c\}$, $\mathrm{ch}(e) = \{j\}$, $\mathrm{ne}(e) = \{f\}$, $\mathrm{bd}(e) = \{c, f\}$, and $\mathrm{ad}(e) = \{c, f, j\}$. The partition $\{a, b, c, d\}$, $\{e, f, g, h\}$, $\{i, j, k\}$ is an example of a chain for G. The sequence of connectivity components $\{a, b, c, d\}$, $\{e, f\}$, $\{g, h\}$, $\{i, j\}$, $\{k\}$ is also a chain for G. Further in G, $\mathrm{an}(e) = \{a, b, c, d, e, f\}$ and $\mathrm{ds}(e) = \{e, f, i, j, k\}$. The set $\mathrm{an}(e)$ is ancestral, but $\mathrm{ds}(e)$ is not an ancestral set.

2.3. *Graphical representation of independency models* A *subway* between nodes u and w in a virtual graph G is a path $\pi : v_1, \dots, v_k$, $k \geq 3$, where $u = v_1 \to v_2$, $\forall\, i = 2, \dots, k-2 \quad v_i - v_{i+1}$, and $v_{k-1} \leftarrow v_k = w$. If no additional edge between the nodes of $T = \{v_1, \dots, v_k\}$ exists in G, it is called a *complex* in G. {Note that Frydenberg [9] used the term 'minimal complex' for what we call a complex and used the word 'complex' for another concept which we consider to be superfluous. We hope that our change simplifies the terminology.} Thus, a complex is nothing but an induced subgraph G_T of the type depicted by Figure 2. The nodes v_1 and v_k are called the *parents* of the complex, the set $\{v_2, \dots, v_{k-1}\}$ the *region* of the complex, and the number $k-2$ is the *degree* of the complex.

The *moral graph* of a CG $G = (N, \mathcal{E})$ denoted by G^{mor} is the virtual graph $(N, \mathcal{E} \cup \mathcal{F})$ where $\mathcal{F} = \{\, u \text{ ---- } v \ ;\ u, v \in N$ such that $\{u, v\} \notin \mathcal{E}$ and there exists a subway between u and v in $G\}$. That is, the type of old edges is saved and the new edges are virts. Note that the moral graph of G can be obtained from G as follows: for each connectivity component C of G all nodes in $\mathrm{pa}_G(C)$ which are not already joined by an edge in G are joined by a virt. {Note that Lauritzen [17] introduced the moral graph as an UG, namely as the underlying graph of our moral graph. This is immaterial from the point of view of the moralization criterion. However, for our proof of equivalence of the separation and moralization criterion we found it very convenient to distinguish the types of edges, and therefore we have introduced the concept of virtual graph and defined the moral graph as a virtual graph.} The term 'moral graph' has a good explanation for the case of DAGs - namely that unmarried parents of every child are married (that is, joined by an edge) in the moral graph. Perhaps it can be applied also in the case of CGs if one interprets connectivity components of a CG as groups of siblings.

Supposing $\langle X, Y|Z\rangle \in T(N)$ and G is a virtual graph over N, we will write $\langle X, Y|Z\rangle_G$ if each route in G from X to Y is hit by Z (equivalently each path). Note the independency model $\{\langle X, Y|Z\rangle \in T(N);\ \langle X, Y|Z\rangle_G\}$ induced by G through this *classic separation* is a graphoid - it suffices to use the result for UGs [22], since for the underlying graph H of the virtual graph G one has $\langle X, Y|Z\rangle_H$ iff $\langle X, Y|Z\rangle_G$.

Let G be a CG, and $\langle X, Y|Z\rangle \in T(N)$. Then we say that $\langle X, Y|Z\rangle$ is *represented* in G according to the *moralization criterion* if every path in the moral graph of $G_{\mathrm{an}(XYZ)}$ from a node of X to a node of Y is hit by Z, and write $\langle X, Y|Z\rangle_G^{mor}$. For every UG G one has $\langle X, Y|Z\rangle_G^{mor}$ iff $\langle X, Y|Z\rangle_G$, i.e. the moralization criterion coincides with the classic criterion for UGs. In case of DAGs it generalizes the moralization criterion from [18].

Thus, every CG G over N defines the model $\{\langle X, Y|Z\rangle \in T(N);\ \langle X, Y|Z\rangle_G^{mor}\}$ over N called the *independency model induced by G*. An independency model over N is then called a *CG-model* if it is induced by some CG over N. Similarly, *UG-models* and *DAG-models* are defined.

Example 2.2. Let us consider the CG G from Figure 1 and ask whether $\langle c, h|a\rangle_G^{mor}$. The ancestral set of $\{a, c, h\}$ is $\{a, b, c, d, g, h\}$. The corresponding

FIG. 3. *Moral graphs $G^{mor}_{\mathrm{an}(ach)}$ and $G^{mor}_{\mathrm{an}(aceh)}$ of the graph from Figure 1.*

induced subgraph is shown in the left picture of Figure 3. It is equal to its moral graph. The only path in that graph from c to h is hit by a and one can conclude that $\langle c,h|a\rangle$ is represented in G according to the moralization criterion. To decide whether $\langle c,h|ae\rangle^{mor}_G$ we construct $G^{mor}_{\mathrm{an}(aceh)}$, which is shown in the right picture of Figure 3. Since the path $c \cdots d \to g - h$ in the moral graph with a virt is not hit by $\{a,e\}$ we conclude that $\neg\langle c,h|ae\rangle^{mor}_G$.

Given a CG G over N, a distribution P over N is called *Markovian with respect to G*, or simply *G-Markovian*, if $\langle X,Y|Z\rangle^{mor}_G$ implies $X \perp\!\!\!\perp Y \mid Z\ [P]$. {Note that some authors [22] in case of UGs and DAGs prefer the terminology 'G is an I-map of P'.} We leave it to the reader to verify:

LEMMA 2.3. *Let G be a CG over N, and H be a subgraph of G (over a set $V \subset N$). Then H is a CG over V and $\langle X,Y|Z\rangle^{mor}_G$ implies $\langle X,Y|Z\rangle^{mor}_H$ for every $\langle X,Y|Z\rangle \in T(V)$. Especially, in the case $V=N$ every H-Markovian probability distribution is also G-Markovian.*

Two CGs G and H over N are *Markov equivalent* if the classes of H-Markovian and G-Markovian probability distributions over N coincide. The importance of the concept of complex is justified by the result from [9], where it is shown that two CGs over N are Markov equivalent iff they have the same underlying graph and the same complexes. In fact, Frydenberg [9] used a weaker definition of Markov equivalence, but as shown in [2], Theorem 3.1 one can also derive the above mentioned characterization from Frydenberg's result.

3. Input list The purport of the notion of local Markov property [18] is that a graphical independency model can be equivalently described by a relatively small list of independency statements, one statement for each node of the graph. As concerns UGs, it was shown in [4] that the graphoid closure of the list of triplets $\langle u, N\setminus \mathrm{ne}(u)u \mid \mathrm{ne}(u)\rangle$, where $u \in N$ and N is the set of nodes of the graph, called the *independency base*, coincides with the class of triplets represented in the UG. In the case of DAGs, the list of triplets corresponding to the local well-numbering Markov property [18] is known as the *causal input list* and it was proved in [28] that a triplet is represented in a DAG iff it belongs to the semigraphoid closure of that list. Note that the semigraphoid closure of the causal input list coincides with its graphoid closure.

In this section we introduce the concept of input list for a CG. This concept corresponds to the (local) chain Markov property [16] (see also [9]) and generalizes the above mentioned analogous concepts for UGs and DAGs. Then we show that a triplet is represented in a CG according to the moralization criterion iff it belongs to the graphoid closure of the introduced input list.

Let G be a CG over N and $\mathcal{B} : B_1, \ldots, B_n$ a chain for G. The *domain* of a node u, written $\mathrm{dom}^{\mathcal{B}}(u)$, is the union of blocks $B_1, \ldots, B_k$, where B_k is the block containing u. The *input list* associated with G and $\mathcal{B}$ is the set of triplets

$$L_G^{\mathcal{B}} = \{ \ \langle u \, , \, \mathrm{dom}^{\mathcal{B}}(u) \setminus \mathrm{bd}_G(u)u \mid \mathrm{bd}_G(u) \, \rangle \ ; \ u \in N \ \}.$$

Note that in the case $\mathrm{dom}^{\mathcal{B}}(u) \setminus \mathrm{bd}_G(u)u = \varnothing$ the corresponding triplet is omitted in $L_G^{\mathcal{B}}$. The triplet $\langle u \, , \, \mathrm{dom}^{\mathcal{B}}(u) \setminus \mathrm{bd}_G(u)u \mid \mathrm{bd}_G(u) \, \rangle$ describes an independency statements interpreted as follows. The variable u is conditionally independent of all its time predecessors and contemporaries $\mathrm{dom}^{\mathcal{B}}(u)$ given its direct causes $\mathrm{pa}_G(u)$ and its symmetric associative influences $\mathrm{ne}_G(u)$. Note, that the input list for a CG generalizes the concepts of the independency base for an UG [4] (if the chain consists of one block) and the concept of the causal input list for a DAG [28] (if the chain consists of singletons).

Example 3.1. Consider once more the graph G from Figure 1 and the chain $\mathcal{B}$: $\{a, b, c, d\}$, $\{e, f\}$, $\{g, h\}$, $\{i, j\}$, $\{k\}$. Then, the domain of a is $\mathrm{dom}^{\mathcal{B}}(a) = \{a, b, c, d\}$ and so are the domains of b, c, and d. Further, $\mathrm{bd}_G(a) = \{b, c\}$, so for node a we have the triplet $\langle a, d|bc\rangle$ in the input list $L_G^{\mathcal{B}}$. For node g, we have $\mathrm{dom}^{\mathcal{B}}(g) = \{a, b, c, d, e, f, g, h\}$ and $\mathrm{bd}_G(g) = \{d, h\}$ so for g we have the triplet $\langle g, abcef|dh\rangle$ in $L_G^{\mathcal{B}}$.

Lemma 3.1. *Supposing G is a CG, the independency model induced by G is a graphoid. Moreover, for every chain $\mathcal{B}$ for G, $\langle X, Y|Z\rangle \in \mathrm{gr}(L_G^{\mathcal{B}})$ implies $\langle X, Y|Z\rangle_G^{mor}$.*

Proof. First, we verify that the CG-model satisfies all graphoid properties.

For symmetry ($\langle X, Y|Z\rangle \longrightarrow \langle Y, X|Z\rangle$) it suffices to realize that the moral graph $H = G^{mor}_{\mathrm{an}(XYZ)}$ for $\langle X, Y|Z\rangle$ coincides with the moral graph for $\langle Y, X|Z\rangle$ as the triplets involve the same variables. Thus, $\langle X, Y|Z\rangle_H$ implies $\langle Y, X|Z\rangle_H$ (graphoid properties for the classic separation - see Subsection 2.3) and therefore $\langle X, Y|Z\rangle_G^{mor}$ implies $\langle Y, X|Z\rangle_G^{mor}$. The same arguments hold for weak union and intersection.

For decomposition ($\langle X, YW|Z\rangle \longrightarrow \langle X, W|Z\rangle$) first realize that the moral graph $K = G^{mor}_{\mathrm{an}(WXZ)}$ for $\langle X, W|Z\rangle$ is a subgraph of the moral graph $H = G^{mor}_{\mathrm{an}(WXYZ)}$ for $\langle X, YW|Z\rangle$. Thus, $\langle X, YW|Z\rangle_H$ implies $\langle X, W|Z\rangle_H$ (see Subsection 2.3) and hence $\langle X, W|Z\rangle_K$.

For contraction ($[\langle X, Y|WZ\rangle \& \langle X, W|Z\rangle] \longrightarrow \langle X, YW|Z\rangle$) we will suppose that $\langle X, Y|WZ\rangle_G^{mor}$ and $\langle X, W|Z\rangle_G^{mor}$. To show $\langle X, YW|Z\rangle_G^{mor}$ it suffices owing to symmetry and intersection (see above) for every $x \in X$ and $w \in YW$ to verify the statement $\langle x, w \,|\, WXYZ \setminus xw\rangle_G^{mor}$. In case $w \in Y$ it follows from $\langle X, Y|WZ\rangle_G^{mor}$ by symmetry and weak union. Thus, consider $w \in W$ and assume, by contradiction, that there exists a path π from $x \in X$ to w in $H = G^{mor}_{\mathrm{an}(WXYZ)}$ that is not hit by $WXYZ \setminus xw$. If all edges of π belong to $K = G^{mor}_{\mathrm{an}(WXZ)}$, then by $\langle X, W|Z\rangle_G^{mor}$ it is hit by Z and thus by $WXYZ \setminus xw$ as well, which contradicts the assumption. Thus, necessarily an edge of π that is not in K must exist. Let us take such an edge $\{u, v\}$, which is closest to x in π, and where u is on the side to x (it may happen that $u = x$). If $\{u, v\}$ is an edge in G, then $v \in \mathrm{an}(Y) \setminus \mathrm{an}(WXZ)$. If $\{u, v\}$ is a virt in H, then there exists a subway $u \to z_1 - \ldots - z_l \leftarrow v$, $l \geq 1$ in G where $z_i \in \mathrm{an}(Y) \setminus \mathrm{an}(WXZ)$ (necessarily for all $i = 1, \ldots, l$). However, in both cases there exists a path in H from x to a node of $z \in \mathrm{an}(Y) \setminus \mathrm{an}(WXZ)$ outside $WXZ \setminus x$. On the other hand, there exists a descending path in G from z to a node $y \in Y$ (necessarily outside $\mathrm{an}(WXZ)$). As this descending path exists in H, one can join the path and obtain a path in H from x to a node in Y outside WZ, which contradicts $\langle X, Y|WZ\rangle_G^{mor}$.

Thus, to verify the second claim of the lemma it suffices to show that each triplet from the input list $L_G^{\mathcal{B}}$ is represented in G according to the moralization criterion. The triplet for $u \in N$ has the corresponding ancestral set $\mathrm{dom}^{\mathcal{B}}(u)$, and moreover $\mathrm{ad}_H(u) = \mathrm{bd}_G(u)$ for the corresponding moral graph H. Hence, $\mathrm{bd}_G(u)$ separates u from the rest of $\mathrm{dom}^{\mathcal{B}}(u)$ in H and the triplet is represented in G according to the moralization criterion. $\square$

Now, we are to prove the converse of Lemma 3.1.

Lemma 3.2. *Let G be a CG and $\mathcal{B}$ a chain for G. Then $\langle X, Y|Z\rangle_G^{mor}$ implies $\langle X, Y|Z\rangle \in \mathrm{gr}(L_G^{\mathcal{B}})$.*

Proof. One can suppose that the considered chain $\mathcal{B}$ has connected blocks. Indeed, each block B_i of the given chain $\mathcal{B}$ can be replaced by a sequence of connectivity components of G, whose union is B_i, and thus a new chain $\mathcal{C}$ with connected blocks is created. As $\mathrm{dom}^{\mathcal{C}}(u) \subset \mathrm{dom}^{\mathcal{B}}(u)$ for every $u \in N$, every triplet $\langle u, \mathrm{dom}^{\mathcal{C}}(u) \setminus \mathrm{bd}_G(u)u \,|\, \mathrm{bd}_G(u)\rangle \in L_G^{\mathcal{C}}$ can be derived by decomposition from its counterpart $\langle u, \mathrm{dom}^{\mathcal{B}}(u) \setminus \mathrm{bd}_G(u)u \,|\, \mathrm{bd}_G(u)\rangle \in L_G^{\mathcal{B}}$. Thus, $L_G^{\mathcal{C}} \subset \mathrm{gr}(L_G^{\mathcal{B}})$ and hence $\mathrm{gr}(L_G^{\mathcal{C}}) \subset \mathrm{gr}(L_G^{\mathcal{B}})$. Therefore we can replace $\mathcal{B}$ by $\mathcal{C}$.

It suffices to show that for every ancestral set $T \subset N$ in the graph G and for $u \in T$, the triplet $\langle u, T \setminus \mathrm{ad}_H(u)u \,|\, \mathrm{ad}_H(u)\rangle$ belongs to $\mathrm{gr}(L_G^{\mathcal{C}})$, where $H = G_T^{mor}$. Indeed, if $\langle X, Y|Z\rangle_G^{mor}$, then by definition $\langle X, Y|Z\rangle_H$, where $H = G_T^{mor}$ for $T = \mathrm{an}(XYZ)$. That is, $\langle X, Y|Z\rangle$ is represented in the underlying graph $\tilde{H}$ of H, and by the already mentioned result for UGs from [4], belongs to the graphoid closure of the independency base of $\tilde{H}$. Thus, to get the desired conclusion $\langle X, Y|Z\rangle \in \mathrm{gr}(L_G^{\mathcal{C}})$ it remains to verify that each triplet from the independency base of $\tilde{H}$, that is of the form $\langle u, T\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle$ for some $u\in T$, belongs to $\mathrm{gr}(L_G^{\mathcal{C}})$.

We prove the claim by induction on the number of blocks of $\mathcal{C}$ which cover T. Thus, supposing $\mathcal{C}: B_1,\dots,B_n$ is a chain for G with connected blocks, we have the following induction hypothesis for every $1\le m\le n$:

Induction hypothesis (I.H.): Let T be an ancestral set with $T\subset B_1\cup\ldots\cup B_m$, $H=G_T^{mor}$, $u\in T$. Then $\langle u, T\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle$ belongs to $\mathrm{gr}(L_G^{\mathcal{C}})$.

Basic step: For $m=1$ observe that $T=B_1$ and $ad_H(u)=\mathrm{bd}_G(u)$ for each $u\in T$. Thus the considered triplet directly belongs to $L_G^{\mathcal{C}}$.

Induction step: Consider $m>1$ and assume that the I.H. holds for $m-1$. Then one can distinguish 3 cases:

(a) $u\in B_m$,
(b) $u\in B_1\cup\ldots\cup B_{m-1}$ and $u\notin \mathrm{pa}_G(B_m)$,
(c) $u\in B_1\cup\ldots\cup B_{m-1}$ and $u\in \mathrm{pa}_G(B_m)$.

In case (a) $ad_H(u)=\mathrm{bd}_G(u)$.Hence $\langle u, B_1\cup\ldots\cup B_m\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle\in L_G^{\mathcal{C}}$ implies by decomposition $\langle u, T\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$.

In case (b) one can suppose $T\cap B_m=B_m$, take $S=T\cap(B_1\cup\ldots\cup B_{m-1})$ which is an ancestral set, and consider the graph $K=G_S^{mor}$. Observe that K is a subgraph of H and $\mathrm{ad}_H(u)=\mathrm{ad}_K(u)$ owing to the assumption of (b). Hence, according to the I.H. applied to S, $\langle u, S\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. Now, for all $w\in B_m$ the case (a) occurs and hence $\langle w, T\setminus \mathrm{ad}_H(w)w \mid \mathrm{ad}_H(w)\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$ implies by weak union and symmetry $\langle u, w\mid T\setminus uw\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. Applying intersection repeatedly gives $\langle u, B_m\mid S\setminus u\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. This together with the claim derived from I.H. implies by contraction $\langle u, T\setminus \mathrm{ad}_H(u)u \mid \mathrm{ad}_H(u)\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$.

Also in case (c) one can suppose $T\cap B_m=B_m$. Then for $w\in B_m\setminus \mathrm{ad}_H(u)$ the case (a) occurs and therefore $\langle w, T\setminus \mathrm{ad}_H(w)w \mid \mathrm{ad}_H(w)\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. Hence by weak union and symmetry $\langle u, w\mid T\setminus uw\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. Moreover, the case (b) holds for $w\in B_1\cup\ldots\cup B_{m-1}\setminus \mathrm{ad}_H(u)u$ (otherwise the edge $\{u,w\}$ in H occurs!) and by analogous procedure derive $\langle u, w\mid T\setminus uw\rangle\in \mathrm{gr}(L_G^{\mathcal{C}})$. Finally, the repeated use of intersection gives the desired conclusion.

Thus, the induction step was verified, which concludes the proof. □

Thus, one can conclude from Lemmas 3.1 and 3.2:

Consequence 3.1. *Let G be a CG and $\mathcal{B}$ a chain for G. Then $\langle X,Y|Z\rangle_G^{mor}$ iff $\langle X,Y|Z\rangle\in \mathrm{gr}(L_G^{\mathcal{B}})$.*

It follows from the preceding consequence that the graphoid closure of the input list does not depend on a particular choice of the input list (that is on the choice of the chain). Moreover, as every independency model induced by a strictly positive distribution is a graphoid (see Subsection 2.1), one can also derive another consequence, which already appeared as Theorem 3.5 in [9].

Consequence 3.2. *Let G be a CG over N, $\mathcal{B}$ a chain for G and P a strictly positive probability distribution over N. Then P is G-Markovian iff P satisfies*

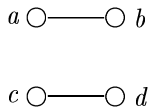

FIG. 4. *Semigraphoid closures depend on the choice of the input list.*

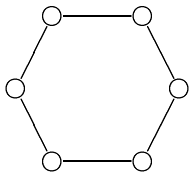

FIG. 5. *An UG-model having a more complex semigraphoid-generating base.*

all the triplets from the input list associated with G and $\mathcal{B}$.

Consequence 3.1 generalizes an analogous result for UGs [4] but also the mentioned result for DAGs [28], if it is understood as a statement about graphoid closure. It can be understood in this way since in the case of DAGs the graphoid and semigraphoid closures of causal input lists coincide. However, this is not true in general, even for UGs as the following example shows.

EXAMPLE 3.2. In case of UGs the semigraphoid closure of the input list may depend on the choice of the chain and may differ from its graphoid closure (that is from the corresponding UG-model). Consider the UG from Figure 4. The chain $\mathcal{B}_1 : \{c,d\}, \{a,b\}$ generates the input list $\{ \langle a, cd|b\rangle, \langle b, cd|a\rangle \}$, whose semigraphoid closure does not contain the triplet $\langle c, ab|d\rangle$. On the other hand, the chain $\mathcal{B}_2 : \{a,b\}, \{c,d\}$ produces the input list $\{ \langle c, ab|d\rangle, \langle d, ab|c\rangle \}$. However, the graphoid closure of both these input lists contains the triplet $\langle ab, cd|\varnothing\rangle$.

REMARK. The reader can object that perhaps we have introduced a wrong concept of input list and perhaps one can associate to every node of a CG another independency statement and perhaps then the semigraphoid closure of the alternative input list can coincide with the CG-model and perhaps the alternative input list can also generalize the causal input list for DAGs. However this is not possible already in the case of UGs. Consider for example the UG from Figure 5. Then one can show that every list of independency statements whose semigraphoid closure is the corresponding UG-model has at least 9 items (we omit an intricate proof of this fact). Thus, there is no hope that one can find a semigraphoid-generating base whose items correspond to the nodes of the UG.

4. Separation criterion We have already mentioned in the Introduction that two equivalent criteria for reading independency statements from a DAG were developed. Namely, the moralization criterion [18] and the separation criterion [10], where one tests whether paths in a DAG are *d-separated*, that is blocked in their nodes (a special definition of blocking is used). The moralization criterion was generalized for CGs [17], but no direct separation criterion for CGs (that is without the auxiliary construction of the moral graph) has been developed so far.

In this section, we introduce the concept of blocking for routes in CGs, called *c-separation* (chain separation), which generalizes the concept of d-separation for routes in DAGs. The corresponding separation criterion for CGs exhibits two main differences from the case of DAGs. First, for CGs one has to consider a wider class of routes (not only paths), for which the blocking is tested (we call them *trails*). Second, the blocking of the route is not defined for nodes of the route, but for its maximal undirected subroutes (we call them *sections*). To formulate the criterion, we have to introduce some special graphical concepts. We show that the blocking of routes is equivalent to the blocking of trails. Then we prove that the introduced separation criterion for CGs is equivalent to the classic moralization criterion.

4.1. *Concept of c-separation* Let G be a CG over N. A *slide* in G from a node u to a node w is a path $v_1, \ldots, v_k$, $k \geq 2$ such that $u = v_1 \to v_2$, $\forall i = 2, \ldots, k-1$ $v_i — v_{i+1}$ and $v_k = w$. The node $u = v_1$ is then called the *top node* of the slide.

By a *section* of a route $\rho : v_1, \ldots, v_k$, $k \geq 1$ in G we mean a maximal undirected subroute $\sigma : v_i — \ldots — v_j$, $1 \leq i \leq j \leq k$ of ρ. Thus, any route decomposes uniquely into sections. The nodes v_i and v_j are then called the *terminals* of the section σ. The node v_i (resp. v_j) is called a *head-terminal* if $i > 1$ and $v_{i-1} \to v_i$ in G (resp. $j < k$ and $v_j \leftarrow v_{j+1}$ in G), otherwise it is called a *tail-terminal* (that is, for v_i, either $i = 1$ or $v_{i-1} \leftarrow v_i$, and for v_j, either $j = k$ or $v_j \to v_{j+1}$). A section σ of a route ρ is called a *head-to-head* section w.r.t. ρ if it has two head-terminals, a *head-to-tail* section w.r.t. ρ if it has one head-terminal and one tail-terminal, and a *tail-to-tail* section w.r.t. ρ if it has two tail-terminals. Note that in the case that the section σ is just one node (i.e. if $i = j$) the node is considered twice as a terminal of σ since two corresponding arrows of ρ should enter the node (each one can be directed either from the node or to the node). This also occurs in the case when $i \neq j$ but $v_i = v_j$. Thus, our classification of sections generalizes an analogous classification of nodes of a route in a DAG [10].

Let σ be a section of a route ρ in G and $Z \subset N$. Then, σ is *blocked* by Z if one of the following cases occurs:

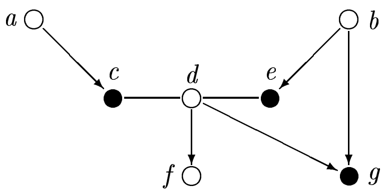

FIG. 6. *One cannot ignore self-intersecting routes in c-separation.*

either σ is a head-to-head section w.r.t. ρ and $ds_G(\sigma) \cap Z = \emptyset$,
or σ is not a head-to-head section w.r.t. ρ, σ is hit by Z, and for at least one of its tail-terminals u, every slide in G to u is hit by Z.

A route in a CG is *c-separated* (chain separated) by Z if at least one of its sections is blocked by Z. Note that a route in a DAG is c-separated iff it is d-separated [22], while a route in an UG is c-separated iff it is hit by Z.

A route which is not c-separated by Z will be called *active* (w.r.t. Z). Let us write the definition explicitly since we will use the concept of active route quite often througout the paper. Thus, a route ρ is active w.r.t. Z iff for every section σ of ρ, one of the following cases occurs:

either σ is a head-to-head section w.r.t. ρ and $ds_G(\sigma) \cap Z \neq \emptyset$,
or σ is a not a head-to-head section w.r.t. ρ and σ is free w.r.t. Z,
or σ is not a head-to-head section w.r.t. ρ and σ is hit by Z, but for each tail-terminal u of σ, there exists a slide in G to u which is free w.r.t. Z.

Having a CG G over N and $\langle X, Y|Z\rangle \in T(N)$, we will say that $\langle X, Y|Z\rangle$ is *represented* in G according to the *separation criterion* if every route in G from a node of X to a node of Y is c-separated by Z and write $\langle X, Y|Z\rangle_G^{sep}$. To understand connection with UGs and DAGs one should realize that although the criteria of separation for UGs, and d-separation for DAGs are usually formulated for paths, they can be equivalently formulated for routes. Thus, for an UG H, one has $\langle X, Y|Z\rangle_H^{sep}$ iff $\langle X, Y|Z\rangle_H$, i.e. the separation criterion generalizes the classic criterion for UGs. Similarly, in the case of DAGs it reduces to the separation criterion for DAGs [10].

However, unlike the case of the separation criteria for UGs and for DAGs the separation criterion for CGs cannot be restricted to paths! The requirement that all paths between two nodes are c-separated is indeed weaker than the requirement that all routes between those nodes are c-separated as the following example shows.

EXAMPLE 4.1. Let us consider the CG G from Figure 6 and ask whether $\langle a, f|ceg\rangle$ is represented in G according to the separation criterion. We observe that the only path in G from a to f is $a \to c — d \to f$. This path is c-separated by ceg as its head-to-tail section $c — d$ is blocked by ceg. However, the route $a \to c — d — e \leftarrow b \to g \leftarrow d \to f$ is not c-separated by ceg since both its head-to-head sections $c — d — e$ and g have a descendant in ceg and all the other sections are free w.r.t. ceg. Therefore, we conclude $\neg\langle a, f|ceg\rangle_G^{sep}$. The reader can check that the moralization criterion gives the same result.

The previous example leads to the question of decidability of the separation criterion for CGs. Since routes allow unlimited repetition of nodes and edges infinitely many routes may exist between two distinct nodes. Fortunately, there exists a reasonable finite class of routes which it is always sufficient to test for c-separation.

A *trail* in a CG G is a route τ in G such that no arrow appears twice in τ, and every section of τ consists of distinct nodes. Note that it follows from the definition that every three consecutive nodes of a trail are distinct. Clearly, every path is a trail. It is evident that in each CG only finitely many trails may exist between two nodes (every CG has finitely many arrows). Moreover, it holds:

LEMMA 4.1. *Let G be a CG over N, $\langle X, Y|Z\rangle \in T(N)$. Then, $\langle X, Y|Z\rangle_G^{sep}$ iff every trail in G from a node of X to a node of Y is c-separated by Z.*

PROOF. It suffices to show that every active route $\rho : v_1, \ldots, v_k$, $k \geq 1$ in G can be shortened to an active trail by a series of changes (we mean active w.r.t. Z). First, we remove multiple occurences of arrows. Thus, suppose that $\{v_i, v_{i+1}\}$, resp. $\{v_j, v_{j+1}\}$, is the first, resp. the last occurence, of the same arrow in ρ (in the order determined by the succession in ρ). Therefore $1 \leq i < j < k$.

In the case of "same orientation", that is $[v_i = v_j$ and $v_{i+1} = v_{j+1}]$, one simply removes the subroute $v_i, \ldots, v_{j-1}$ from ρ. As all remaining sections are unchanged by this removal, they are not blocked in the shortened route.

In the case of "contrary orientation", that is $[v_i = v_{j+1}$ and $v_{i+1} = v_j]$, one removes the subroute $v_i, \ldots, v_j$ from ρ. Thus, by pooling of the sections $\sigma_1 : v_m, \ldots, v_i$, $1 \leq m \leq i$ and $\sigma_2 : v_{j+1}, \ldots, v_n$, $j + 1 \leq n \leq k$ of the original route the section $\sigma_3 : v_m, \ldots, v_i = v_{j+1}, \ldots, v_n$ of the shortened route is created. As the other sections were untouched, it suffices to verify that σ_3 is not blocked.

If σ_3 is a head-to-head section and $v_i \leftarrow v_{i+1}$, then σ_1 is a head-to-head section (w.r.t. the original route) and $\mathrm{ds}_G(\sigma_1) = \mathrm{ds}_G(\sigma_3)$. As σ_1 was not blocked in the original route $\mathrm{ds}_G(\sigma_1) \cap Z \neq \varnothing$ and therefore σ_3 is not blocked in the shortened route.

If σ_3 is a head-to-head section and $v_i \to v_{i+1}$, then one can find within the removed subroute $v_i, \ldots, v_j$ a descending route from v_i to a head-to-head section σ_4 (of the original route). As $\mathrm{ds}_G(\sigma_4) \subset \mathrm{ds}_G(\sigma_3)$ and σ_4 is not blocked, the section σ_3 will not be blocked.

In case σ_3 is not a head-to-head section it suffices to show for every tail-terminal u of σ_3 that either there exists a free slide to u, or σ_3 is free. Without loss of generality we suppose $u = v_m$ (otherwise, in case $u = v_n$ use the same consideration where we replace v_m by v_n and σ_1 by σ_2). Thus, we distinguish the following subcases:

- If σ_1 is hit, then there exists a free slide to $u = v_m$ since v_m is a tail-terminal of σ_1.
- If σ_1 is free and $v_i \leftarrow v_{i+1}$, then $v_{i+1} \notin Z$ (otherwise the section of the original route which follows after σ_1 is blocked) and the route $v_{i+1} \rightarrow v_i — \ldots — v_m$ can be shortened to a free slide to $u = v_m$.
- If σ_1 is free, $v_i \rightarrow v_{i+1}$ (that is $v_{j+1} \rightarrow v_j$) and σ_2 is hit, then $v_i = v_{j+1}$ is a tail-terminal of σ_2, and therefore there exists a free slide from a node w to v_i. This slide can be lengthened by σ_1 to a free route from w to v_m. The obtained route can be shortened to a free slide.
- If both σ_1 and σ_2 are free, then σ_3 is free.

Thus, an active route without multiple occurences of arrows was obtained. Clearly, each of its sections now can be shortened to consist of distinct nodes. As this modification does not change the types of sections and their terminals, the shortened route, more precisely trail, is active. □

EXAMPLE 4.2. Consider once more the CG G from Figure 1. To find out whether $\langle c, h|a\rangle_G^{sep}$ consider a trail from c to h. If the trail from c starts by the line $c — a$, then its starting tail-to-tail section $c — a — b — d$ meets a and therefore is blocked by a (there is no slide to tail-terminals of the section). If the trail starts by $c \rightarrow e$, then it contains a head-to-head section with no descendant in a (either $c \rightarrow e — f \leftarrow d$ or $f \rightarrow k \leftarrow g$), and is blocked as well. So, we conclude that $\langle c, h|a\rangle_G^{sep}$. To see that $\neg\langle c, h|ae\rangle_G^{sep}$ realize that the path $c \rightarrow e — f \leftarrow d \rightarrow g — h$ is active w.r.t. ae. The section $e — f$ is a head-to-head section that has a descendant in ae, namely e. The other sections are not head-to-head sections and are not hit by ae. So, no section is blocked by ae. Note that the mentioned path is the only active trail from c to h (w.r.t ae).

4.2. *Equivalence of separation and moralization criteria*

LEMMA 4.2. *Let G be a CG over N, $\langle X, Y|Z\rangle \in T(N)$. Then $\langle X, Y|Z\rangle_G^{sep}$ implies $\langle X, Y|Z\rangle_G^{mor}$.*

PROOF. We show $\neg\langle X, Y|Z\rangle_G^{mor} \Rightarrow \neg\langle X, Y|Z\rangle_G^{sep}$, that is, if there exists a path π in $G^{mor}_{\mathrm{an}(XYZ)}$ from X to Y which is outside Z, then there exists an active route (w.r.t. Z) in G between X and Y. The original path π will be gradually transformed into the desired route in G.

Whenever $\{u, v\}$ is a virt of π, then by the definition of the moral graph, there exists a subway between u and v in $G_{\mathrm{an}(XYZ)}$ and the virt can be replaced by that subway. Thus, the original path π is transformed into a route ρ in $G_{\mathrm{an}(XYZ)}$ such that all sections of ρ, with possible exception of head-to-head sections, are free (w.r.t. Z).

The second step is to transform ρ into an active route in $G_{\mathrm{an}(XYZ)}$ between X and Y. More exactly, one can transform ρ into a route with no blocked head-to-head sections, whose other sections are free. Indeed, if σ is a blocked head-to-head section of ρ, then $\sigma \subset \mathrm{an}_G(XYZ)$ says $\mathrm{ds}_G(\sigma) \cap XYZ \neq \varnothing$ and owing to $\mathrm{ds}_G(\sigma) \cap Z = \varnothing$ there exists a free descending path from σ to $w \in XY$. If $w \in X$ (resp. $w \in Y$), then one can attach this descending path from σ to w to a subroute of ρ from Y (resp. from X) to σ and obtain a route between X and Y which has fewer blocked head-to-head sections than ρ and whose other sections are free w.r.t. Z. □

LEMMA 4.3. *Let G be a CG over N, $\langle X, Y|Z\rangle \in T(N)$. Then $\langle X, Y|Z\rangle_G^{mor}$ implies $\langle X, Y|Z\rangle_G^{sep}$.*

PROOF. We show $\neg\langle X, Y|Z\rangle_G^{sep} \Rightarrow \neg\langle X, Y|Z\rangle_G^{mor}$, that is if there exists an active route τ in G from X to Y, then there exists a free path (w.r.t. Z) in $G^{mor}_{\mathrm{an}(XYZ)}$ from X to Y. The route τ will be gradually transformed into the desired path.

First, we realize that τ belongs to $G_{\mathrm{an}(XYZ)}$. Indeed, if $\tau \setminus \mathrm{an}(XYZ) \neq \varnothing$, then one can find a head-to-head section σ of τ such that $\mathrm{ds}(\sigma) \cap \mathrm{an}(XYZ) = \varnothing$ (if a section of τ outside $\mathrm{an}(XYZ)$ is not a head-to-head section, the outgoing arrow of τ leads to another section outside $\mathrm{an}(XYZ)$). But such a section is blocked, which contradicts the assumption.

We transform τ into a free route (w.r.t. Z) in $G^{mor}_{\mathrm{an}(XYZ)}$, namely for every hit section σ of τ we replace a certain subroute $\bar{\sigma}$ of τ, containing σ, by a free detour (path) in $G^{mor}_{\mathrm{an}(XYZ)}$. The modification depends on the type of σ.

- If σ is a head-to-head section, then the corresponding subroute has the form $\bar{\sigma} : \bar{u} \to u - \ldots - v \leftarrow \bar{v}$. Clearly, $\bar{u}, \bar{v} \notin Z$, as otherwise a neighbouring section of τ is blocked. If $\bar{u} = \bar{v}$, then no detour is necessary, otherwise $\{\bar{u}, \bar{v}\}$ is an edge in $G^{mor}_{\mathrm{an}(XYZ)}$ and $\bar{\sigma}$ is replaced by that edge.
- If σ is a head-to-tail section, the subroute is $\bar{\sigma} : \bar{u} \to u - \ldots - v$, where v is the tail-terminal of σ. Evidently $\bar{u}, v \notin Z$. Moreover, there exists a free slide $w_1 \to w_2 - \ldots - w_l = v$, $l \geq 2$ in $G_{\mathrm{an}(XYZ)}$. In case $\bar{u} = w_1$ the detour is made of that slide, otherwise the edge $\{\bar{u}, w_1\}$ in $G^{mor}_{\mathrm{an}(XYZ)}$ is added to that slide.
- If σ is a tail-to-tail section, then the corresponding subroute is just σ. Evidently, for each of both tail-terminals of σ there exists a free slide to the tail-terminal. The detour will consist of these two slides and of the edge in $G^{mor}_{\mathrm{an}(XYZ)}$ between their top nodes (unless the top nodes coincide).

The obtained free route ρ in $G^{mor}_{\mathrm{an}(XYZ)}$ can be shortened to a path by removing all subroutes between two possible different occurences of the same node. □

So, we can summarize Lemmas 4.2 and 4.3 as follows.

CONSEQUENCE 4.1. *Let G be a CG over N, and $\langle X, Y|Z\rangle \in T(N)$. Then $\langle X, Y|Z\rangle_G^{mor}$ iff $\langle X, Y|Z\rangle_G^{sep}$.*

REMARK. In fact, there exists another surprisingly simple formulation of the c-separation criterion which is much more similar to the d-separation criterion. Namely that $\langle X, Y|Z\rangle_G^{sep}$ iff for every route ρ in G there exists a section σ of ρ such that

either σ is a head-to-head section w.r.t. ρ and σ is outside Z
or σ is not a head-to-head section w.r.t. ρ and σ is hit by Z.

To see that this is an equivalent formulation of c-separation realize that whenever one has an active route ρ from X to Y then it can be modified into a *superactive route* that is such a route ϱ that for every section σ of ϱ, the section σ is hit by Z iff σ is a head-to-head section w.r.t. ϱ. Indeed, if there exists a head-to-head section σ in ρ which is not hit by Z then there exists a descending path π in G from a node $u \in \sigma$ to a node $v \in Z$ which is outside Z with exception of v. Thus, one can insert a new 'artificial' subroute into ρ in place of occurence of u within σ: the first part of the subroute copies π in direction from u to v, the second part copies π in the opposite direction. A similar trick can be done with hit non-head-to-head sections of ρ with help of free slides to their tail-terminals. Therefore any active route ρ can be patched to become a superactive route ϱ.

There is a hitch in the above mentioned formulation of the c-separation: a potentially infinite number of routes. Nevertheless, perhaps it is a good argument against a common objection that c-separation is too complicated with comparison with d-separation. What is however probably more important is that this peculiar formulation could become a basis of an efficient computer implementation of c-separation.

5. Completeness of the separation criterion In this section, we prove the completeness of the separation criterion for CGs, namely that a triplet $\langle X, Y|Z\rangle$ is represented in a CG whenever it is a valid independency statement in every Markovian distribution. As CGs were developed to describe CI-structures for strictly positive probability distributions, we prove the completeness with respect to the class of these distributions.

5.1. *Outline of the method* General features of the construction can be outlined as follows. We prove for a given CG G and a triplet $\langle X, Y|Z\rangle$ which is not represented in the graph (according to the separation criterion) that there exists a G-Markovian distribution that does not satisfy the triplet. Special properties of CG-models will allow us to consider an elementary triplet $\langle a, b|Z\rangle$, where $a \in X, b \in Y$, instead of $\langle X, Y|Z\rangle$.

To show that there exists a Markovian distribution which does not satisfy $\langle a, b|Z\rangle$, we perform two basic steps:

- instead of G, we consider a subgraph H of G (over the same set of nodes) such that $\langle a, b|Z\rangle$ is not represented in H, and
- we construct a Markovian distribution with respect to H (and therefore Markovian with respect to G) which does not satisfy $\langle a, b|Z\rangle$.

The aim of the first step is clear: the simpler the graph is, the simpler the corresponding construction of the distribution may be. The choice of the subgraph in our method generalizes an analogous method used in the completeness proof for UGs and DAGs [12, 11], where the subgraph simply is an active path w.r.t. Z possibly supplemented by some additional edges to ensure the path is also active in the subgraph. This is also the basic idea of the choice of the subgraph in our method: we take an active trail (w.r.t. Z) that has minimal complexity in a certain sense and add several suitable additional edges. Now the reader may understand why the c-separation criterion is necessary for this type of consideration. However, in the case of CGs the situation is much more complicated as, in general, one cannot avoid 'self-intersection' of the trail and also the additional edges can unavoidably cause 'cycles' (in the underlying graph).

As concerns the second basic step (that is, the proper construction) we construct a binary Markovian distribution as a 'product' of special 'two-variable' potentials which correspond to the edges of the (moral) graph of H. Our potentials are strictly positive and are parameterized by numbers from the interval $(-1, 1)$. The main problem is to choose the parameters to ensure that $\langle a, b|Z\rangle$ does not hold in the obtained distribution. As a tool for verification we will introduce a certain numerical characteristic (a function) for every binary probability distribution, denoted by κ, with an important property: an independency statement occurs if and only if the function κ has certain values zero. Then we decompose the considered trail between a and b into simple segments and show that κ is nonzero for the whole trail iff it is nonzero for each segment. The question of whether κ is nonzero for a segment will appear to be a local problem, that is, vanishing of κ for a segment will depend only on the parameters of potentials for edges related to the segment. Thus, for each type of a segment one can derive by simple calculation that κ is nonzero iff all the parameters for related edges are nonzero. Therefore, every choice of nonzero parameters will solve our problem.

5.2. *The choice of a subgraph* In this first subsection we formulate all properties of the chosen subgraph which will be used in subsequent subsections to justify our arguments. For this purpose we need some additional concepts.

5.2.1. *Additional graphical concepts* Throughout this subsection G is a CG over N, $Z \subset N$ a set of its nodes, and $\rho : v_1, \ldots, v_m$, $m \geq 1$, a (fixed) route in G which is active w.r.t. Z.

Suppose that $\alpha : v_i — v_{i+1} — \ldots — v_j$, $1 \leq i \leq j \leq m$, is a hit section of ρ. By a *free zone* of α we mean a maximal free subroute of α, that is $\zeta :$ $v_k, \ldots, v_l$, $i \leq k \leq l \leq j$, such that $[v_s \notin Z$ for $k \leq s \leq l]$, $[v_{k-1} \in Z$ or $k = i]$ and $[v_{l+1} \in Z$ or $l = j]$. A free zone is a *terminal free zone* if it contains v_i or v_j, otherwise it is a *middle free zone*.

The remaining special definitions are actually applied only under the condition that every section of ρ consists of distinct nodes (for example in the case that ρ is a trail).

Suppose that $\varepsilon : u_1 \to u_2 — \ldots — u_r$, $r \geq 2$ is a slide in G. In case u_r is a

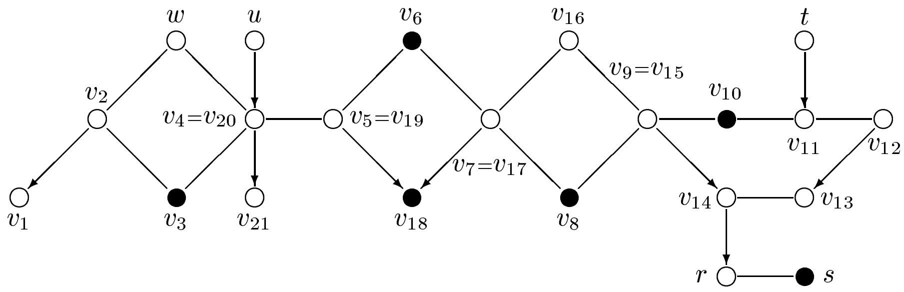

FIG. 7. *An example illustrating additional graphical concepts.*

terminal of a section α (of ρ) we say that α is a *host section* for ε. {Later, under a special situation (= a special route ρ), host sections for slides will be uniquely determined. Then, we will write *the* host section for ε instead of a host section for ε.} We say that a slide ε is *regular* with respect to its host section α if there exists a node u_k, $2 \leq k \leq r$ of ε such that $\{u_1, \ldots, u_{k-1}\}$ does not meet α and $u_k, \ldots, u_r$ copies the section α. {More concretely, if $\alpha : v_i, \ldots, v_j\,, 1 \leq i \leq j \leq m$, and $u_r = v_i$ (resp. $u_r = v_j$), then $r-k \leq j-i$ and $u_{r-s} = v_{i+s}$ (resp. $u_{r-s} = v_{j-s}$) for $s = 0, \ldots, r-k$.} The node u_k is then called the *heel* of ε.

A section $\alpha : v_i - \ldots - v_j$, $1 \leq i \leq j \leq m$ of ρ meets *regularly* another section β of ρ if there exist $i \leq k \leq l \leq j$ such that $\{v_i, \ldots, v_{k-1}\} \cup \{v_{l+1}, \ldots, v_j\}$ does not meet β and $v_k, \ldots, v_l$ copies the section β. {More concretely, in case $\beta : v_t, \ldots, v_n\,, 1 \leq t \leq n \leq m$ there exists $t \leq r \leq n - l + k$ such that either $v_{r+s} = v_{k+s}$ for $s = 0, \ldots, l-k$, or $v_{r+s} = v_{l-s}$ for $s = 0, \ldots, l-k$. Note that symmetrically $\{v_t, \ldots, v_{r-1}\} \cup \{v_{r+l-k+1}, \ldots, v_n\}$ does not meet α since β consists of distinct nodes.}

Let α be a section of ρ. By a *downroute* for α we mean a descending path $\delta : u_1, \ldots, u_r\,, r \geq 2$, from a node of α to a node in Z such that $[u_s \notin Z$ for $1 \leq s \leq r-1]$ and $[u_s \notin \alpha$ for $2 \leq s \leq r]$. The node u_1 is then the *top node* of δ and α is called a *host section* for δ. {Again, under a special later situation, host sections for downroutes will be uniquely determined and we will write *the* host section for δ instead of a host section for δ.}

EXAMPLE 5.1. To illustrate the concepts above let us consider the CG from Figure 7, the trail $\rho : v_1, \ldots, v_{21}$ and the set $Z = \{v_3, v_6, v_8, v_{10}, v_{18}, s\}$. Clearly, ρ is active w.r.t. Z. Its hit head-to-head section v_{18} has no free zone. On the other hand, the hit tail-to-tail section $\alpha : v_2, \ldots, v_{12}$ has two terminal free zones, namely v_2 and v_{11}, v_{12}, and three middle free zones, namely v_4, v_5 and v_7 and v_9. The slide $u \to v_4 - w - v_2$ is an example of a non-regular slide to the tail-terminal v_2 of α. The slide $\varepsilon : t \to v_{11} - v_{12}$ is a regular slide to the other tail-terminal v_{12} of α. The top node of ε is t, the heel of ε is v_{11}. The sections $v_2, \ldots, v_{12}$ and v_{15}, v_{16}, v_{17} do not meet regularly, while $v_2, \ldots, v_{12}$ and v_{19}, v_{20} meet regularly. The free head-to-head section v_{13}, v_{14} has a downroute $v_{14} \to r - s$ whose top node is v_{14}.

5.2.2. *Properties of the chosen subgraph* All properties of the subgraph that we need are collected in the following summary. Since the proof is too long and technical, it is shifted to the Appendix (Subsections 6.2 - 6.5).

Summary 5.1. *Let G be a CG over N and $\langle a, b|Z\rangle \in E(N)$ such that $\neg\langle a, b|Z\rangle_G^{sep}$. Then one can find*

- *an active trail τ between a and b, and*
- *a free slide to every tail-terminal of every hit section of τ, and*
- *a downroute for every free head-to-head section of τ,*

such that the following conditions hold.

(a) Every downroute does not meet another downroute, any slide or any section of τ except its host section in its top node.
(b) Every above mentioned free slide ε to a tail-terminal of a hit section σ of τ is regular with respect to σ. Moreover, ε does not meet another slide or a downroute. It can meet at most one other section except its host section σ, but only in its top node. The other section is then a hit section and the shared top node belongs to a middle free zone of the other section.
(c) Every pair of sections which share a node meet regularly. Moreover, if this occurs, then one of the sections is a free tail-to-tail section, the other one is a hit section and the shared part belongs to a middle free zone of the hit section. {Evidently, a free section cannot meet two or more free zones since the sections meet regularly.} Every free tail-to-tail section can meet at most one other section.
(d) Every hit section can meet several top nodes of free slides and several free tail-to-tail sections, but every one of its middle free zones can meet at most one top node or at most one free tail-to-tail section, and cannot meet both.

Moreover, the subgraph H over N consisting of the edges of the trail, the slides and the downroutes mentioned above has the following properties. {To help the reader to have a good insight let us mention that it follows from the previous facts that every node can belong to at most two of the mentioned objects, that is the sections (of the trail), the slides and the downroutes.}

(e) $\neg\langle a, b|Z\rangle_H^{sep}$.
(f) H has no undirected cycles. Equivalently, for every connectivity component C of H the induced subgraph H_C is a tree.
(g) Every connectivity component of H has at most 2 arrows heading to it in H (that is, directed into a node of the connectivity component) and if it happens, then the arrows originate from different nodes which are not joined by an edge in H.

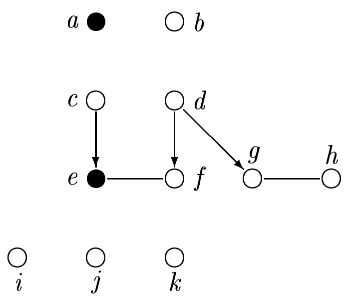

FIG. 8. *An example of the chosen subgraph.*

(h) For every pair of distinct nodes $u, v \in N$ at most one complex in H having u and v as its parents can exist.

EXAMPLE 5.2. To illustrate the procedure described in Summary 5.1 let us consider the graph G from Figure 1 and the elementary triplet $\langle c, h|ae\rangle \in E(N)$, which is not represented in G according to the separation criterion (see Example 4.2). The only active trail from c to h (w.r.t. ae) is the path $c \to e — f \leftarrow d \to g — h$; therefore it has to be chosen as the trail in Summary 5.1. As the only hit section $e — f$ of that path does not have a tail-terminal and no free head-to-head section occurs in the path, no free slides and no downroutes are chosen in the procedure described in Summary 5.1. Thus, the obtained subgraph is exactly the mentioned path - see Figure 8.

5.2.3. *Decomposition of a trail into segments* Suppose that we have already chosen the subgraph H for a triplet $\langle a, b|Z\rangle \in E(N)$, consisting of the trail, downroutes and slides as described above. The trail is naturally partitioned by the tail-terminals of its sections (here we take all sections, both hit and free). Note, for clarity, that the end nodes of the trail a and b are also tail-terminals. Nevertheless, we will consider an even finer partition of the trail.

Under the situation described by Summary 5.1 let $\tau : v_1, \ldots, v_m$, $m \geq 2$, be the chosen active trail and S the set of tail-terminals of τ and heels of regular slides, described there. A *segment* of τ is its subroute $v_i, \ldots, v_j$ such that $1 \leq i < j \leq n$, $v_i, v_j \in S$ and $[v_s \notin S$ for $i < s < j]$. Thus, the set S decomposes the trail uniquely into segments $[u_i, u_{i+1}]$, where $a = u_1, \ldots, u_n = b$, $n \geq 2$, is the corresponding sequence of nodes of S. It is no problem to check that every segment belongs exclusively to one of the following 7 types:

I. the segment contains a free head-to-tail section,
II. the segment contains a free (non-singleton) tail-to-tail section,
III. the segment contains a hit head-to-head section,

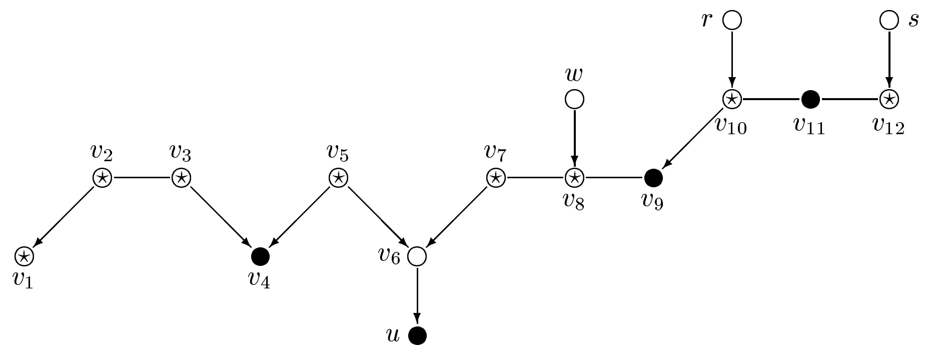

FIG. 9. *Segments of a trail.*

IV. the segment contains a free head-to-head section,
V. the segment is the shared part of a slide and its host (hit) section,
VI. the segment contains a hit part of a head-to-tail section,
VI. the segment contains a hit part of a tail-to-tail section.

EXAMPLE 5.3. To illustrate the procedure of decomposition of the chosen trail into segments let us consider the graph from Figure 9, the trail $v_1,\ldots,v_{12}$ and the set $Z = \{v_4, v_9, v_{11}, u\}$. The trail has tail-terminals $v_1, v_2, v_3, v_5, v_7, v_{10}, v_{12}$ and heels v_8, v_{10}, v_{12}. As in the last section v_{10}, v_{11}, v_{12} the tail-terminals coincide with the heels, the corresponding sequence of tail-terminals and heels is $u_1 = v_1$, $u_2 = v_2$, $u_3 = v_3$, $u_4 = v_5$, $u_5 = v_7$, $u_6 = v_8$, $u_7 = v_{10}$, $u_8 = v_{12}$. Every node of this sequence is indicated by an asterisk in the picture. It gives a decomposition into 7 segments: The segment $[u_1, u_2] = [v_1, v_2]$ contains a free head-to-tail section v_1 and represents type **I.**, the segment $[u_2, u_3] = [v_2, v_3]$ contains a free tail-to-tail section v_2, v_3 and represents type **II.**, the segment $[u_3, u_4] = [v_3, v_5]$ contains a hit head-to-head section v_4 and represents type **III.**, the segment $[u_4, u_5] = [v_5, v_7]$ contains a free head-to-head section v_6 and represents type **IV.**, the segment $[u_5, u_6] = [v_7, v_8]$ is the common part of the hit head-to-tail section v_7, v_8, v_9 and the free slide w, v_8, v_7 to its tail-terminal and represents type **V.**, the segment $[u_6, u_7] = [v_8, v_{10}]$ contains the hit part of the head-to-tail section v_7, v_8, v_9 and represents type **VI.**, the segment $[u_7, u_8] = [v_{10}, v_{12}]$ contains the hit part of the tail-to-tail section v_{10}, v_{11}, v_{12} and represents type **VII.**

The following lemma describes an important property of the decomposition into segments. Its proof is too technical (one has to distinguish many cases) to be presented here. Subsection 6.6 of the Appendix contains instructions as to

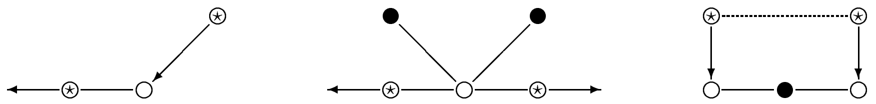

FIG. 10. *Illustrative pictures for Observations 5.1, 5.2 and 5.3.*

how a precise proof can be carried out.

LEMMA 5.1. *Suppose that H is the chosen subgraph for $\langle a,b|Z\rangle \in E(N)$ from Summary 5.1 and $a = u_1,\ldots,u_n = b$ is the corresponding sequence of tail-terminals and heels of chosen free slides. Then, under the assumption $n \geq 3$, for each $i = 1,\ldots,n-2$ every path in H^{mor} from $u_1 \ldots u_i$ to $u_{i+2} \ldots u_n$ is hit by $u_{i+1}Z$. In particular, $\forall\ i = 1,\ldots,n-2 \quad \langle u_i, u_n|u_{i+1}Z\rangle_H^{mor}$.*

Supposing that $\langle u,v|Z\rangle \in E(N)$ and K is a virtual graph over N, by an *essential* path for $[u,v]$ in K (w.r.t. Z) we mean a path in K between u and v that is outside Z. By an *associated* path for $[u,v]$ in K (w.r.t. Z) we mean a path in K from a node $w \in N$ to an inner node t of an essential path for $[u,v]$ (that is $u \neq t \neq v$) that is outside Z with the possible exception of w. The following series of observations analyzes essential and associated paths in H^{mor} for every type of segment. They will be utilized later and their proofs are left to the reader: it is a mechanical application of facts from Summary 5.1 already used in the proof of Lemma 5.1 described in the Appendix.

OBSERVATION 5.1. *If $[u,v]$ is a segment of type* **I.**, *the only essential path for $[u,v]$ is the considered segment and no associated path for $[u,v]$ exists (see the left picture of Figure 10).*

OBSERVATION 5.2. *If $[u,v]$ is a segment of type* **II.**, *that is, a non-singleton free tail-to-tail section, the only essential path for $[u,v]$ is the considered segment. An associated path for $[u,v]$ can occur only when a middle free zone of a hit section meets the free tail-to-tail section and consists of nodes of the (unique) free zone (see the middle picture of Figure 10).*

OBSERVATION 5.3. *If $[u,v]$ is a segment of type* **III.**, *the only essential path for $[u,v]$ is the virt between u and v and no associated path for $[u,v]$ exists (see Figure 10, the right picture).*

OBSERVATION 5.4. *If $[u,v]$ is a segment of type* **IV.**, *then two essential paths for $[u,v]$ exist. The first one is the virt between u and v, the second one is the*

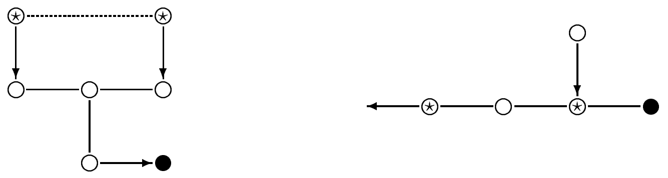

FIG. 11. *Illustrative pictures for Observations 5.4 and 5.5.*

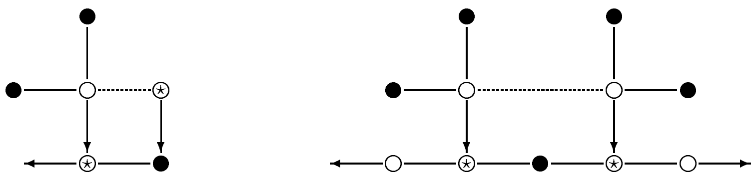

FIG. 12. *Illustrative pictures for Observations 5.6 and 5.7.*

segment. Every associated path for $[u,v]$ is a part of the corresponding downroute (see the left picture of Figure 11).

OBSERVATION 5.5. *If $[u,v]$ is a segment of type* **V.***, the only essential path for $[u,v]$ is the considered segment and no associated path for $[u,v]$ exists (see Figure 11, the right picture).*

OBSERVATION 5.6. *If $[u,v]$ is a segment of type* **VI.** *(for example u is a heel of a slide, and v is the contiguous tail-terminal of a neighbouring section), then the only essential path for $[u,v]$ consists of the part of the slide between u and its top node w, and of the virt between w and v. An associated path for $[u,v]$ can be made of nodes of a possible middle free zone of another hit section which meets w (see the left picture of the Figure 12).*

OBSERVATION 5.7. *If $[u,v]$ is a segment of type* **VII.** *(that is u and v are heels of slides), then the only essential path for $[u,v]$ is made of the parts of slides between their heels and top nodes, and of the virt between those top nodes. Associated paths for $[u,v]$ can occur and consist of nodes of possible middle free zone(s) of other hit section(s) which meet the top node(s) (see Figure 12, the right picture).*

5.3. *Construction of a Markovian distribution* In this subsection we describe how to construct a Markovian distribution with respect to the chosen subgraph H. First, we introduce a numerical characteristic of independence, denoted by κ, and derive its basic properties (Subsubsection 5.3.1). Then, we define special elementary potentials and formulate their basic properties (Subsubsection 5.3.2). They form the basis of the proper construction of the distribution which follows (Subsubsection 5.3.3). Decomposition of the trail into segments described in Subsubsection 5.2.3 enables us to convert the crucial problem of verification that the constructed distribution does not satisfy the considered elementary triplet into a series of local questions, that is, questions depending only on small sets of nodes, closely related to the segments (Subsubsection 5.3.4). An analysis of each of the 7 possible types of segment then makes it possible to prove the main lemma saying that a nonzero choice of parameters solves our central problem (Subsubsection 5.3.5).

5.3.1. *Numerical characteristic of independence* We will extensively use a numerical characteristic of independence represented by the function κ. Suppose that $\langle a,b|Z\rangle \in E(N)$, $\mathbf{x}_Z \in \{0,1\}^Z$ and f is a potential on $\{0,1\}^W$, where $abZ \subset W \subset N$. Then, $\kappa : N \times N \times \{0,1\}^Z \times (\{0,1\}^W \to \Re) \to \Re$ is the function defined by

$$\kappa(a,b|\mathbf{x}_Z\|f) \;=\; f^{abZ}(00\mathbf{x}_Z)\cdot f^{abZ}(11\mathbf{x}_Z) - f^{abZ}(01\mathbf{x}_Z)\cdot f^{abZ}(10\mathbf{x}_Z)\,.$$

Note that the definition applies also in case $Z=\varnothing$ when it is a number denoted by $\kappa(a,b|-\|f)$. The value of κ depends only on the marginal for f on abZ and therefore f can always be replaced by its marginal involving variables abZ:

$$\kappa(a,b|\mathbf{x}_Z\|f) \;=\; \kappa(a,b|\mathbf{x}_Z\|f^X) \qquad \text{for } abZ \subset X \subset W\,.$$

The values of κ allow us to characterize conditional independence relationships in the case of binary probability distributions.

Lemma 5.2. *Let P be a probability distribution on $\{0,1\}^N$ and $\langle a,b|Z\rangle \in E(N)$. Then $a \perp\!\!\!\perp b\,|\,Z\,[P]$ iff $[\forall \mathbf{x}_Z \in \{0,1\}^Z \quad \kappa(a,b|\mathbf{x}_Z\|P)=0\,]$.*

Proof. For necessity realize that $P^Z(\mathbf{x}_Z)=0$ implies $P^{abZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_Z)=0$ for all $\mathbf{x}_a,\mathbf{x}_b \in \{0,1\}$ and hence $\kappa(a,b|\mathbf{x}_Z\|P)=0$. However, in case $P^Z(\mathbf{x}_Z)>0$ owing to $a \perp\!\!\!\perp b\,|\,Z\,[P]$ one can write

$$P^{abZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_Z) \;=\; \frac{P^{aZ}(\mathbf{x}_a\mathbf{x}_Z)\cdot P^{bZ}(\mathbf{x}_b\mathbf{x}_Z)}{P^Z(\mathbf{x}_Z)}$$

for $\mathbf{x}_a,\mathbf{x}_b \in \{0,1\}$, $\mathbf{x}_Z \in \{0,1\}^Z$. It remains to substitute the formula into the definition of $\kappa(a,b|\mathbf{x}_Z\|P)$. For sufficiency we realize that the vanishing of $\kappa(a,b|\mathbf{x}_Z\|P)$ implies

$$P^{abZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_Z)\cdot P^{abZ}(\mathbf{y}_a\mathbf{y}_b\mathbf{x}_Z) \;=\; P^{abZ}(\mathbf{x}_a\mathbf{y}_b\mathbf{x}_Z)\cdot P^{abZ}(\mathbf{y}_a\mathbf{x}_b\mathbf{x}_Z)\,.$$

for $\mathbf{x}_a,\mathbf{x}_b,\mathbf{y}_a,\mathbf{y}_b \in \{0,1\}$, $\mathbf{x}_Z \in \{0,1\}^Z$. The summation in this formula through $\mathbf{y}_a,\mathbf{y}_b \in \{0,1\}$ yields $a \perp\!\!\!\perp b\,|\,Z\,[P]$. $\square$

It may be difficult to verify directly that $\kappa(a,b|\mathbf{x}_Z\|P)$ is nonzero. The following lemma enables us to transform this into a couple of 'simpler' analogous questions.

LEMMA 5.3. *Let P be a distribution on $\{0,1\}^N$; $Z \subset N$; $a,b,c \in N \setminus Z$ distinct. If $a \perp\!\!\!\perp b \,|\, cZ\ [P]$, then one has for every $\mathbf{x}_Z \in \{0,1\}^Z$:*

$$\kappa(a,b|\mathbf{x}_Z\|P) \neq 0 \quad \textit{iff} \quad [\,\kappa(a,c|\mathbf{x}_Z\|P) \neq 0 \ \&\ \kappa(c,b|\mathbf{x}_Z\|P) \neq 0\,].$$

PROOF. First, we verify for every $\mathbf{x}_Z \in \{0,1\}^Z$ the formula

$$\kappa(a,b|\mathbf{x}_Z\|P) \cdot P^{cZ}(0\mathbf{x}_Z) \cdot P^{cZ}(1\mathbf{x}_Z) = \kappa(a,c|\mathbf{x}_Z\|P) \cdot \kappa(c,b|\mathbf{x}_Z\|P)\ .$$

Indeed, substitute the equalities $P^{abZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_Z) = P^{abcZ}(\mathbf{x}_a\mathbf{x}_b0\mathbf{x}_Z) + P^{abcZ}(\mathbf{x}_a\mathbf{x}_b1\mathbf{x}_Z)$ into the definition of $\kappa(a,b|\mathbf{x}_Z)$ (we omit the symbol of P within $\kappa(\cdot,\cdot|\cdot\|\cdot)$ in the rest of the proof) and obtain by easy calculation:

$$\begin{aligned}\kappa(a,b|\mathbf{x}_Z) \ =\ & \sum \{\,\kappa(a,b|\mathbf{x}_c\mathbf{x}_Z)\ ;\ \mathbf{x}_c = 0,1\,\}\\ &+\ P^{abcZ}(000\mathbf{x}_Z)\cdot P^{abcZ}(111\mathbf{x}_Z) \ +\ P^{abcZ}(001\mathbf{x}_Z)\cdot P^{abcZ}(110\mathbf{x}_Z)\\ &-\ P^{abcZ}(010\mathbf{x}_Z)\cdot P^{abcZ}(101\mathbf{x}_Z) \ -\ P^{abcZ}(011\mathbf{x}_Z)\cdot P^{abcZ}(100\mathbf{x}_Z)\,.\end{aligned}$$

Owing to $a \perp\!\!\!\perp b \,|\, cZ\ [P]$ and Lemma 5.2 the terms $\kappa(a,b|\mathbf{x}_c\mathbf{x}_Z)$ vanish and we can write:

$$\begin{aligned}&\kappa(a,b|\mathbf{x}_Z)\cdot P^{cZ}(0\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z) \ =\\ &\quad +\ \{P^{abcZ}(000\mathbf{x}_Z)\cdot P^{cZ}(0\mathbf{x}_Z)\}\cdot\{P^{abcZ}(111\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z)\}\\ &\quad +\ \{P^{abcZ}(001\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z)\}\cdot\{P^{abcZ}(110\mathbf{x}_Z)\cdot P^{cZ}(0\mathbf{x}_Z)\}\\ &\quad -\ \{P^{abcZ}(010\mathbf{x}_Z)\cdot P^{cZ}(0\mathbf{x}_Z)\}\cdot\{P^{abcZ}(101\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z)\}\\ &\quad -\ \{P^{abcZ}(011\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z)\}\cdot\{P^{abcZ}(100\mathbf{x}_Z)\cdot P^{cZ}(0\mathbf{x}_Z)\}\,.\end{aligned}$$

Owing to $a \perp\!\!\!\perp b \,|\, cZ\ [P]$ the identity

$$P^{abcZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_c\mathbf{x}_Z)\cdot P^{cZ}(\mathbf{x}_c\mathbf{x}_Z) \ =\ P^{acZ}(\mathbf{x}_a\mathbf{x}_c\mathbf{x}_Z)\cdot P^{bcZ}(\mathbf{x}_b\mathbf{x}_c\mathbf{x}_Z)$$

can be substituted in every expression in curly brackets. We leave it to the reader to verify that the obtained result is nothing but the expanded expression $\kappa(a,c|\mathbf{x}_Z)\cdot\kappa(c,b|\mathbf{x}_Z)$ which concludes the proof of the formula. The formula implies the claim of the lemma except in the case $P^{cZ}(0\mathbf{x}_Z)\cdot P^{cZ}(1\mathbf{x}_Z) = 0$. However, in the case $P^{cZ}(0\mathbf{x}_Z) = 0$ one can suppose that $P^Z(\mathbf{x}_Z) > 0$ (otherwise the claim of the lemma is trivial) and write

$$P^{abZ}(\mathbf{x}_a\mathbf{x}_b\mathbf{x}_Z) \ =\ P^{abcZ}(\mathbf{x}_a\mathbf{x}_b1\mathbf{x}_Z) = \frac{P^{acZ}(\mathbf{x}_a1\mathbf{x}_Z)\cdot P^{bcZ}(\mathbf{x}_b1\mathbf{x}_Z)}{P^{cZ}(1\mathbf{x}_Z)}$$

by $a \perp\!\!\!\perp b \,|\, cZ\ [P]$. This gives $\kappa(a,b|\mathbf{x}_Z) = 0$ and the claim of the lemma is also verified in this case. Analogous justification holds in case $P^{cZ}(1\mathbf{x}_Z) = 0$. □

5.3.2. *Elementary potentials* The distribution that we are going to define is based on special potentials associated with pairs of adjacent nodes in a graph. For every $u, v \in N$ $u \neq v$ and every parameter $\gamma \in (-1,1)$ we define a potential on the set $\{0,1\}^{\{u,v\}}$, called an *elementary potential* and denoted by R^{uv}_{γ} as follows:

$$\begin{aligned} R^{uv}_{\gamma} \,:\, (0,0) &\rightarrow (1+\gamma)/2 \\ (0,1) &\rightarrow (1-\gamma)/2 \\ (1,0) &\rightarrow (1-\gamma)/2 \\ (1,1) &\rightarrow (1+\gamma)/2\,. \end{aligned}$$

We leave it to the reader to verify its properties mentioned in the following lemma.

LEMMA 5.4. *The elementary potential R^{uv}_{γ} is a strictly positive function on $\{0,1\}^{\{u,v\}}$ which is constant only in the case $\gamma = 0$. The following formulas relate marginalization, multiplication and inverse.*

(a) $(R^{uv}_{\gamma})^{\{u\}} \equiv 1 \equiv (R^{uv}_{\gamma})^{\{v\}}$ *for $u, v \in N$ distinct, $\gamma \in (-1,1)$*
(b) $(R^{uw}_{\gamma} \cdot R^{wv}_{\delta})^{\{u,v\}} = R^{uv}_{\gamma\cdot\delta}$ *for $u, v, w \in N$ distinct, $\gamma, \delta \in (-1,1)$*
(c) $(R^{uv}_{\gamma})^{-1} = \frac{4}{(1+\gamma)\cdot(1-\gamma)} \cdot R^{uv}_{-\gamma}$ *for $u, v \in N$ distinct, $\gamma \in (-1,1)$.*

We will later utilize substantially the fact that if a function f is a special product of elementary potentials a simple formula for $\kappa(\cdot,\cdot|\cdot\|f)$ holds. These properties are illustrated by Figure 13 where a Greek letter accompanying an edge represents the parameter of the elementary potential for the pair of variables at the end the edge.

LEMMA 5.5. *The following formulas hold.*

(a) *If $f = R^{uv}_{\alpha}$ where $\alpha \in (-1,1)$, $Z = \varnothing$ (see Fig. 13v), then $\kappa(u,v|-\|f) = \alpha$.*
(b) *If $f = R^{uw}_{\alpha} \cdot R^{vw}_{\beta} \cdot R^{wz}_{\gamma}$ where $\alpha, \beta, \gamma \in (-1,1)$ and $Z = \{z\}$ (see Fig. 13i), then $\kappa(u,v|\mathbf{x}_Z\|f) = \alpha\cdot\beta\cdot(1-\gamma^2)/4$ for every $\mathbf{x}_Z \in \{0,1\}^Z$.*
(c) *If $f = R^{uw}_{\alpha} \cdot R^{vw}_{\beta} \cdot R^{wz_1}_{\gamma} \cdot R^{wz_2}_{\delta}$ where $\alpha, \beta, \gamma, \delta \in (-1,1)$ and $Z = \{z_1, z_2\}$ (see Fig. 13ii), then $\kappa(u,v|\mathbf{x}_Z\|f) = \alpha\cdot\beta\cdot(1-\gamma^2)\cdot(1-\delta^2)/16$ for every $\mathbf{x}_Z \in \{0,1\}^Z$.*
(d) *If $f = R^{ut}_{\alpha}\cdot R^{tw}_{\beta}\cdot R^{vw}_{\gamma}\cdot R^{tz_1}_{\delta}\cdot R^{wz_2}_{\varepsilon}$ where $\alpha, \beta, \gamma, \delta, \varepsilon \in (-1,1)$ and $Z = \{z_1, z_2\}$ (see Fig. 13iii), then $\kappa(u,v|\mathbf{x}_Z\|f) = \alpha\cdot\beta\cdot\gamma\cdot(1-\delta^2)\cdot(1-\varepsilon^2)/16$ for every $\mathbf{x}_Z \in \{0,1\}^Z$.*
(e) *If $f = R^{uv}_{\alpha}\cdot R^{uw}_{\beta}\cdot R^{vw}_{\gamma}\cdot R^{wz}_{\delta}$ where $\beta, \gamma, \delta \in (-1,1)$, $\alpha = -\beta\cdot\gamma$ and $Z = \{z\}$ (see Fig. 13iv), then $\kappa(u,v|\mathbf{x}_Z\|f) = -\beta\cdot(1-\beta^2)\cdot\gamma\cdot(1-\gamma^2)\cdot\delta^2/16$ for every $\mathbf{x}_Z \in \{0,1\}^Z$.*
(f) *If $f = R^{ut}_{\alpha}\cdot R^{tw}_{\beta}\cdot R^{vw}_{\gamma}\cdot R^{tz_1}_{\delta}\cdot R^{tz_2}_{\varepsilon}\cdot R^{wz_3}_{\zeta}\cdot R^{wz_4}_{\eta}$ where $\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta \in (-1,1)$ and $Z = \{z_1, z_2, z_3, z_4\}$ (see Fig. 13vi), then*

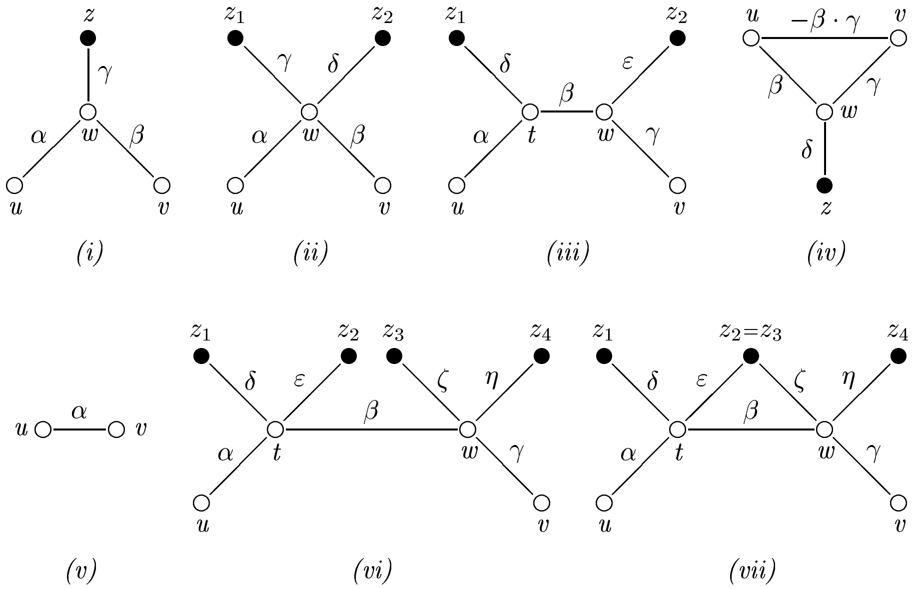

FIG. 13. *Pictures illustrating Lemma 5.5.*

$\kappa(u,v|\mathbf{x}_Z\|f) \;=\; \alpha\cdot\beta\cdot\gamma\cdot(1-\delta^2)\cdot(1-\varepsilon^2)\cdot(1-\zeta^2)\cdot(1-\eta^2)/256$
for every $\mathbf{x}_Z \in \{0,1\}^Z$.

(g) *The same result as in the preceding case holds with modification* $z_2 = z_3$ *(see Fig. 13vii).*

PROOF. We leave it to the reader to verify the formulas. However, we give two important hints. First, to avoid difficulties the reader is advised to write and expand the expression for $\kappa(u,v|\mathbf{x}_Z\|f)$ as a polynomial whose primitive terms are $(1+\alpha)$, $(1-\alpha)$, $(1+\beta)$ etc. Then some terms will be cancelled. After that cancellation one can expand the remaining expression in α, β etc.

The second hint is that one need not verify the formula for all possible values of $\mathbf{x}_Z$ but just for one possible configuration, say for configuration 0_Z. The reason is the form of elementary potentials. For example, in the case (b), the expression for $\kappa(u,v|1\|f)$ written in α,β,γ can be conceived as the expression for $\kappa(u,v|0\|f)$ in which γ is replaced by $(-\gamma)$. But in the resulting formula for $\kappa(u,v|0\|f)$ the parameter γ is squared, and therefore $\kappa(u,v|1\|f) = \kappa(u,v|0\|f)$! A similar trick can be used in the other cases: parameters of the edges which enter Z can have alternative signs but the result is always the same. □

5.3.3. *Core construction of the distribution* Let us suppose that the subgraph H satisfying the conditions of Summary 5.1 for a triplet $\langle a,b|Z\rangle \in E(N)$ was

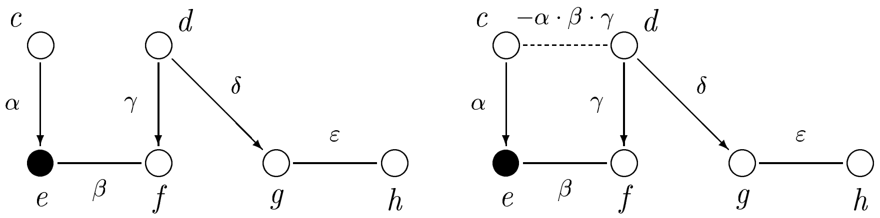

FIG. 14. *Construction of a Markovian distribution.*

already fixed and choose for every edge $\{u,v\}$ in H a parameter $\gamma(u,v)$ from the interval $(-1,1)$. For a connectivity component C of H, $\mathcal{H}(C)$ denotes the class of lines in C and arrows heading to C. Under these circumstances we define a probability distribution P on $\{0,1\}^N$ as follows:

1. Define a strictly positive potential Q_C on $\{0,1\}^{C\cup \mathrm{pa}(C)}$ as the product of elementary potentials for edges in $\mathcal{H}(C)$:

$$Q_C(\mathbf{x}_{C\cup \mathrm{pa}(C)}) \;=\; \prod_{\{u,v\}\in\mathcal{H}(C)} R^{uv}_{\gamma(u,v)}(\mathbf{x}_{uv}) \qquad \text{for } \mathbf{x}_{C\cup \mathrm{pa}(C)} \in \{0,1\}^{C\cup \mathrm{pa}(C)}\,.$$

   Of course, in case $\mathcal{H}(C)=\varnothing$ the product over the empty set is 1 and therefore $Q_C \equiv 1$.
2. For every connectivity component C of H we compute from Q_C the conditional probability $Q_{C|\,\mathrm{pa}(C)}$ on C given $\mathrm{pa}(C)$:

$$Q_{C|\,\mathrm{pa}(C)}(\mathbf{x}_C|\mathbf{x}_{\mathrm{pa}(C)}) \;=\; \frac{Q_C(\mathbf{x}_{C\cup \mathrm{pa}(C)})}{Q_C^{\mathrm{pa}(C)}(\mathbf{x}_{\mathrm{pa}(C)})} \qquad \text{for } \mathbf{x}_{C\cup \mathrm{pa}(C)} \in \{0,1\}^{C\cup \mathrm{pa}(C)}\,.$$

3. Define the distribution P as the product of these conditional probabilities:

$$P(\mathbf{x}_N) \;=\; \prod_{C\in\mathcal{C}} Q_{C|\,\mathrm{pa}(C)}\,(\mathbf{x}_C|\mathbf{x}_{\mathrm{pa}(C)}) \qquad \text{for } \mathbf{x}_N \in \{0,1\}^N\,,$$

   where $\mathcal{C}$ denotes the class of all connectivity components of H.

As Q_C is strictly positive, its marginal $Q_C^{\mathrm{pa}(C)}$ is also strictly positive and we are allowed to divide by it. Evidently, P is a strictly positive binary probability distribution over N.

EXAMPLE 5.4. To illustrate the construction let us continue with Example 5.2. Each edge of the chosen subraph H can be endowed with a parameter from the interval $(-1,1)$. Namely α for $c\to e$, β for $e - f$, γ for $d\to f$, δ for $d\to g$ and ε for $g - h$ - see the left picture of Figure 14, where remaining nodes from Figure 8 are omitted. Now, we can express the potentials for connectivity components of

H (see the first step of the construction): $Q_{ef} = R^{ce}_{\alpha} \cdot R^{ef}_{\beta} \cdot R^{df}_{\gamma}$, $Q_{gh} = R^{dg}_{\delta} \cdot R^{gh}_{\varepsilon}$ and $Q_u \equiv 1$ for any remaining node u of H. Using Lemma 5.4(a)(b) one can compute their marginals on parent sets: $Q^{cd}_{ef} = R^{cd}_{\alpha\cdot\beta\cdot\gamma}$, $Q^{d}_{gh} \equiv 1$ and $Q^{\varnothing}_{u} \equiv 2$ for remaining $u \in N$. Therefore the conditional probabilities from the second step are: $Q_{ef|cd} = R^{ce}_{\alpha} \cdot R^{ef}_{\beta} \cdot R^{df}_{\gamma} / R^{cd}_{\alpha\cdot\beta\cdot\gamma}$, $Q_{gh|d} = R^{dg}_{\delta} \cdot R^{gh}_{\varepsilon}$, and $Q_{u|\varnothing} = 1/2$ for any other $u \in N$. Hence, the corresponding probability distribution has the following form (H has seven singleton connectivity components):

$$P = 2^{-7} \cdot R^{ce}_{\alpha} \cdot R^{ef}_{\beta} \cdot R^{df}_{\gamma} \cdot R^{dg}_{\delta} \cdot R^{gh}_{\varepsilon} / R^{cd}_{\alpha\cdot\beta\cdot\gamma} \,.$$

.

LEMMA 5.6. *The probability distribution P constructed as described above is H-Markovian for any choice of parameters.*

PROOF. One can use Theorem 4.1 from [9]. The condition (ii) there says that a (strictly positive) distribution is H-Markovian iff it is the product of conditional probabilities $\prod_{C\in\mathcal{C}} Q_{C|\,\mathrm{pa}(C)}$, where $Q_{C|\,\mathrm{pa}(C)}$ is computed from an $H^{mor}_{C\cup\mathrm{pa}(C)}$-Markovian distribution over $C \cup \mathrm{pa}(C)$. However, our potential Q_C is nothing but a multiple of such a distribution (by a strictly positive constant) since Q_C by definition factorizes according to the graph $H_{C\cup\mathrm{pa}(C)}$ and therefore according to the graph $H^{mor}_{C\cup\mathrm{pa}(C)}$ (this factorization is equivalent to the Markov property for UGs by Proposition 1 from [18]). □

The following lemma gives a more suitable formula for the constructed probability distribution.

LEMMA 5.7. *For every virt u --- v in H^{mor} there exists a unique complex λ in H having u, v as its parents. Supposing H and parameters for its edges are fixed we introduce parameters also for virts in H^{mor} by the formula*

$$\gamma(u,v) \;=\; -\prod_{\{t,w\}\in\mathcal{L}} \gamma(t,w)\,,$$

where $\mathcal{L}$ is the set of edges of the complex λ in H which has u and v as its parents. Then the probability distribution P defined above has the form:

$$P(\mathbf{x}_N) \;=\; c \cdot \prod_{\{u,v\}\in\mathcal{H}^*} R^{uv}_{\gamma(u,v)}(\mathbf{x}_{uv}) \qquad \text{for } \;\mathbf{x}_N \in \{0,1\}^N \,,$$

where $c > 0$ is a strictly positive normalizing constant (depending on the graph H and on the parameters) and $\mathcal{H}^*$ is the set of edges of H^{mor}.

PROOF. Suppose that u, v are different parents of a connectivity component C of H (for example nodes of a virt in H^{mor}). Summary 5.1(g) says that C can have at most two parents. Moreover, it implies that there exist a unique $s_1 \in C$ with $u \to s_1$ in H, and a unique $s_l \in C$ with $v \to s_l$ in H, and by Summary 5.1(f) with Lemma 2.1(ii) a unique undirected path $s_1, \ldots, s_l$, $l \geq 1$ in C between them (of course, it may happen that $s_1 = s_l$ and the path has only one node). To verify that the subway $\lambda : u \rightarrow s_1 — \ldots — s_l \leftarrow v$ is a complex in H realize that:

a) no other edges can appear among $\{s_1, \ldots, s_l\}$ as otherwise they have to be lines and an undirected cycle exists in C which contradicts Summary 5.1(f),
b) no additional edge between u (resp. v) and $\{s_1, \ldots, s_l\}$ exists as then three arrows heading to C exist, which contradicts Summary 5.1(g),
c) no edge between u and v exists, which also follows from Summary 5.1(g).

The uniqueness of the complex follows from Summary 5.1(h). Moreover, $u \dashrightarrow v$ in H^{mor} and we have shown that there is a one-to-one correspondence between virts in H^{mor} and two-parent connectivity components of H.
Now, Lemma 2.1(iii) implies that $s_1, \ldots, s_l$ can be prolonged into a sequence $s_1, \ldots, s_k$, $k \geq l$ of all nodes of C, such that every s_i, $i > 1$ is joined by a line with just of its one predecessors. This enables us to derive a simple form of the desired marginal for the potential $Q_C = \prod_{\{u,v\} \in \mathcal{H}(C)} R^{uv}_{\gamma(u,v)}$ (which is over the set $\{u, v, s_1, \ldots, s_k\}$). First, we sum through s_k, as only one edge in $\mathcal{H}(C)$ enters s_k in case $k > l$ by Lemma 5.4(a) the result is also a product of elementary potentials, but the term associated with edge entering s_k disappears. Hence by induction $Q^{\lambda}_C = \prod_{\{u,v\} \in \mathcal{L}} R^{uv}_{\gamma(u,v)}$, where $\mathcal{L}$ is the set of edges of the complex λ. By repeated application of Lemma 5.4(b) we get $Q^{\{u,v\}}_C = R^{uv}_{\delta}$, where $\delta = \prod_{\{u,v\} \in \mathcal{L}} \gamma(u,v)$ and hence by Lemma 5.4(c) $(Q^{\{u,v\}}_C)^{-1} = t \cdot R^{uv}_{-\delta} = t \cdot R^{uv}_{\gamma(u,v)}$ where $t = \frac{4}{(1-\delta)\cdot(1+\delta)}$ is strictly positive.
A similar procedure can be used to show that $Q^u_C \equiv 1$ for every component C of H with $\mathrm{pa}(C) = \{u\}$, and $Q^{\varnothing}_C = 2$ and for any component C of H with $\mathrm{pa}(C) = \varnothing$.
To verify the formula we realize that every edge of H belongs to $\mathcal{H}(C)$ for uniquely determined component C. Similarly, an assignment of virts in H^{mor} to two-parent connectivity components of H was established above. Thus, we can write:

$$\begin{aligned}
P &= \prod_{C \in \mathcal{C}} Q_C \cdot (Q^{\mathrm{pa}(C)}_C)^{-1} \\
&= \left(\prod_{C \in \mathcal{C}} \prod_{\{t,w\} \in \mathcal{H}(C)} R^{tw}_{\gamma(t,w)}\right) \cdot \left(\prod_{C \in \mathcal{C}} (Q^{\mathrm{pa}(C)}_C)^{-1}\right) \\
&= \prod_{\{t,w\} \in \mathcal{H}} R^{tw}_{\gamma(t,w)} \cdot c \cdot \prod_{\{u,v\} \in \mathcal{H}^* \setminus \mathcal{H}} R^{uv}_{\gamma(u,v)} ,
\end{aligned}$$

where $c > 0$ is a constant. $\square$

Example 5.5. To illustrate Lemma 5.7 let us consider the graph in the left picture of Figure 14. The only virt in H^{mor} is $c \dashrightarrow d$. The corresponding complex in H is $c \rightarrow e — f \leftarrow d$ with parameters α, β, γ. Thus, $-\alpha \cdot \beta \cdot \gamma$ is the parameter for $c \dashrightarrow d$. The right picture of Figure 14 shows H^{mor} with corresponding parameters (the remaining nodes of the original graph from Figure

8 are omitted). The formula from Example 5.4 then implies by Lemma 5.4(c):

$$P = 2^{-5} \cdot \frac{1}{(1+\alpha\cdot\beta\cdot\gamma)\cdot(1-\alpha\cdot\beta\cdot\gamma)} \cdot R^{ce}_{\alpha} \cdot R^{ef}_{\beta} \cdot R^{df}_{\gamma} \cdot R^{dg}_{\delta} \cdot R^{gh}_{\varepsilon} \cdot R^{cd}_{-\alpha\cdot\beta\cdot\gamma} \,.$$

5.3.4. *Localization* In Subsubsection 5.2.3 we showed how to decompose the considered trail into segments. This allows us to break down the central problem of showing that the constructed distribution does not satisfy the considered elementary triplet $\langle a,b|Z\rangle$ for an appropriate choice of parameters:

CONSEQUENCE 5.1. *Suppose that P is any of the binary probability distributions constructed for H in Subsubsection 5.3.3 and $a = u_1,\ldots,u_n = b$ is the sequence from Subsubsection 5.2.3. If $n \geq 3$, then $u_i \perp\!\!\!\perp u_n \,|\, u_{i+1} Z\ [P]$ for every $i = 1,\ldots,n-2$. Moreover, $\neg\{\, a \perp\!\!\!\perp b \,|\, Z\ [P]\,\}$ iff*

$$[\,\exists\ \mathbf{x}_Z \in \{0,1\}^Z \ \textit{such that}\ \forall\ i = 1,\ldots,n-1 \quad \kappa(u_i,u_{i+1}|\mathbf{x}_Z\|P) \neq 0\,]\,.$$

PROOF. Lemma 5.6 says that P is H-Markovian. Thus $u_i \perp\!\!\!\perp u_n \,|\, u_{i+1} Z\ [P]$ by Lemma 5.1. The second claim follows from Lemma 5.2 by a possibly repeated application of Lemma 5.3. □

Thus, it remains to ensure that $\kappa(u,v|\mathbf{x}_Z\|P) \neq 0$ for each segment $[u,v]$. To show this we will utilize the idea of 'localization' to a certain small set of nodes D described in the following lemma.

LEMMA 5.8. *Let $K = (N,\mathcal{K})$ be a virtual graph and P a probability distribution on $\{0,1\}^N$ of the form $P = c \cdot \prod_{\{t,w\}\in\mathcal{K}} R^{tw}_{\gamma(t,w)}$, where $c > 0$ is a normalizing constant and $\gamma(t,w) \in (-1,1)$ a parameter for each $\{t,w\} \in \mathcal{K}$. Suppose that $\langle u,v|Z\rangle \in E(N)$ is given such that there exists a path in K from u to v outside Z. Let us denote by D the set of nodes which belongs to an essential path for $[u,v]$ (see Subsubsection 5.2.3 for definition) in K w.r.t. Z or to an associated path for $[u,v]$ in K w.r.t. Z, and $\mathcal{D}$ denotes the class of edges of all these paths. Let us put:*

$$d(\mathbf{x}_D) \;=\; \prod_{\{t,w\}\in\mathcal{D}} R^{tw}_{\gamma(t,w)}(\mathbf{x}_{tw}) \qquad \textit{for } \mathbf{x}_D \in \{0,1\}^D \,.$$

Then, for every $\mathbf{x}_Z \in \{0,1\}^Z$, $\kappa(u,v|\mathbf{x}_Z\|P) \neq 0$ iff $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d) \neq 0$.

PROOF. Evidently, $u,v \in D$. Moreover,

(5.1) if $t \in D \setminus Zuv$ and $\{t,w\} \in \mathcal{K}$ for some $w \in N$, then $\{t,w\} \in \mathcal{D}$.

Let us introduce the set E of nodes $t \in N$, such that there exists a path from u to t made of edges of $\mathcal{K} \setminus \mathcal{D}$ which is outside Z with possible exception of t. $\mathcal{E}$ denotes the class of edges of $\mathcal{K} \setminus \mathcal{D}$ in E. Then $v \notin E$, as otherwise there exists a path between u and v outside Z made of edges of $\mathcal{K} \setminus \mathcal{D}$ which contradicts the definition of $\mathcal{D}$. Using (5.1) we can conclude $D \cap E \subset Zu$. Hence, one can derive using the definition of E:

$$\text{(5.2)} \qquad \text{if } t \in E \setminus Zu \text{ and } \{t,w\} \in \mathcal{K} \text{ for some } w \in N\,, \text{ then } \{t,w\} \in \mathcal{E}\,.$$

Finally, we put $F = (N \setminus DE) \cup (DE \cap Zv)$ and $\mathcal{F} = \mathcal{K} \setminus (\mathcal{D} \cup \mathcal{E})$. Then clearly $N = DEF$, $D \cap F \subset Zv$, $u \notin F$, $E \cap F \subset Z$ and $\mathcal{K}$ decomposes into $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$. Further, using (5.1),(5.2) and the definition of E one derives that $\bigcup_{\{t,w\}\in\mathcal{F}} \{t,w\} \subset F$ and moreover by $DE \cap F \subset Zv$:

$$\text{(5.3)} \qquad \text{if } t \in F \setminus Zv \text{ and } \{t,w\} \in \mathcal{K} \text{ for some } w \in N\,, \text{ then } \{t,w\} \in \mathcal{F}\,.$$

Thus, one can write $P = c \cdot \prod_{\{t,w\}\in\mathcal{K}} R^{tw}_{\gamma(t,w)} = c \cdot d \cdot e \cdot f$, where $e(\mathbf{x}_E) = \prod_{\{t,w\}\in\mathcal{E}} R^{tw}_{\gamma(t,w)}(\mathbf{x}_{tw})$ is a strictly positive potential over E and $f(\mathbf{x}_F) = \prod_{\{t,w\}\in\mathcal{F}} R^{tw}_{\gamma(t,w)}(\mathbf{x}_{tw})$ can be considered as a strictly positive potential over F. We put here $e \equiv 1$ in case $\mathcal{E} = \varnothing$ and $f \equiv 1$ in case $\mathcal{F} = \varnothing$. The process of marginalization of P to uvZ can be done separately: owing to (5.2) summing through variables in $E \setminus Zu$ concerns only e, owing to (5.3) summing through variables in $F \setminus Zv$ concerns only f and owing to (5.1) summing through variables in $D \setminus Zuv$ concerns only d. Hence

$$P^{uvZ} = c \cdot d^{uvZ\cap D} \cdot e^{uZ\cap E} \cdot f^{vZ\cap F}$$

and one can write:

$$\begin{aligned} \kappa(u,v|\mathbf{x}_Z\|P) \;=\;& c^2 \cdot e^{uZ\cap E}(0_u\mathbf{x}_{Z\cap E}) \cdot e^{uZ\cap E}(1_u\mathbf{x}_{Z\cap E}) \\ & \cdot f^{vZ\cap F}(0_v\mathbf{x}_{Z\cap F}) \cdot f^{vZ\cap F}(1_v\mathbf{x}_{Z\cap F}) \cdot \kappa(u,v|\mathbf{x}_{Z\cap D}\|d^{uvZ\cap D})\,. \end{aligned}$$

Since all terms in $\kappa(u,v|\mathbf{x}_Z\|P)$ except $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d^{uvZ\cap D}) = \kappa(u,v|\mathbf{x}_{Z\cap D}\|d)$ are strictly positive the claim of the lemma is evident. □

5.3.5. *Main lemma* Now, we are able to prove directly the main lemma of the section.

LEMMA 5.9. *Suppose that $H = (N,\mathcal{H})$ is the chosen subgraph for $\langle a,b|Z\rangle \in E(N)$ from Summary 5.1 and P any of the constructed probability distributions for H from Subsubsection* 5.3.3. *Then*

$$\neg\{\, a \perp\!\!\!\perp b \,|\, Z\,[P]\,\} \qquad \textit{iff} \quad [\,\forall\,\{t,w\} \in \mathcal{H}\;\; \gamma(t,w) \neq 0\,]\,.$$

PROOF. Let us consider the decomposition $a = u_1, \ldots, u_n = b$ of the considered trail into segments described in Subsubsection 5.2.3. By Consequence 5.1 it suffices to show for some $\mathbf{x}_Z \in \{0,1\}^Z$ (in fact we show it for arbitrary such $\mathbf{x}_Z$) that $[\kappa(u,v|\mathbf{x}_Z\|P) \neq 0$ for every segment $[u,v]\,]$ iff $[\gamma(t,w) \neq 0$ for every $\{t,w\} \in \mathcal{H}]$. This can be shown with help of Lemma 5.8 where $K = H^{mor}$: owing to Lemma 5.7 its assumptions are fulfilled for the constructed distribution P and the question as to when $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d) \neq 0$ for the corresponding function d remains to be answered.

This can be done separately for each of 7 possible types of sectors mentioned in Subsubsection 5.2.3. In each case the corresponding class of edges $\mathcal{D}$ has a special form and application of Lemma 5.4 leads to one of the formulas mentioned in Lemma 5.5 which will imply that $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d) \neq 0$ iff parameters of certain class of edges $\mathcal{L}_{[u,v]} \subset \mathcal{H}$ are nonzero. We leave to the reader to check that every edge in $\mathcal{H}$ belongs to some $\mathcal{L}_{[u,v]}$.

I. If $[u,v]$ contains a free head-to-tail section, then by Observation 5.1 D consists of the nodes of the segment only. By consecutive application of Lemma 5.4(b) $d^{\{u,v\}} = R_{\delta}^{uv}$ where δ is the product of parameters of edges of the segment. Lemma 5.5(a) says that $\kappa(u,v|-\|d) \neq 0$ iff $\delta \neq 0$. Thus, $\mathcal{L}_{[u,v]}$ consists of the edges of the segment.

II. If $[u,v]$ contains a free tail-to-tail section, then by Observation 5.2 the set D is either the mere segment $[u,v]$ or the segment with one branch originating from an inner node of the segment or the segment with two branches which possibly may originate from the same inner node. The application Lemma 5.4(b) leads to one of the formulas from Lemma 5.5(a)-(d). The formulas imply that $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d) \neq 0$ iff all the parameters associated with edges of the segments are nonzero. Thus, $\mathcal{L}_{[u,v]}$ again consists of the edges of the segment.

III. If $[u,v]$ contains a hit head-to-head section, then one applies Observation 5.3. Then $D = \{u,v\}$ and by Lemma 5.5(a) $\kappa(u,v|\mathbf{x}_{Z\cap D}\|d) \neq 0$ iff $\gamma(u,v) \neq 0$. That means (see Lemma 5.7) that $\mathcal{L}_{[u,v]}$ is the class of the edges of the segment.

IV. If $[u,v]$ contains a free head-to-head section, then one uses Observation 5.4: $\mathcal{D}$ consists of the virt, edges of the segment and edges of the downroute. Lemma 5.4(b) then leads to the situation from Lemma 5.5(e) (owing to Lemma 5.7). Thus, $\mathcal{L}_{[u,v]}$ consists both of the edges of the segment and the edges of the downroute.

V. If $[u,v]$ is the shared part of a slide and its host hit section, then apply Observation 5.5, Lemma 5.4(b), Lemma 5.5(a) to get the conclusion that $\mathcal{L}_{[u,v]}$ consists of the edges of $[u,v]$.

VI. If $[u,v]$ contains a hit part of a head-to-tail section (for example u is the heel of a slide and v is a tail-terminal of a neighbouring section), then Observation 5.6, Lemma 5.4(b), Lemma 5.5(a),(c) imply that $\mathcal{L}_{[u,v]}$ consists of the nodes of the segment and of the nodes of the slide between its top node and its heel.

VII. If $[u,v]$ contains a hit part of a tail-to-tail section, then use Observation 5.7 and Lemma 5.4(b) leads to application of one of formulas from Lemma 5.5(a),(c),(f),(g). In any of these cases one concludes that $\mathcal{L}_{[u,v]}$ consists both of the edges of the segment and of the edges of slides between their heels and top nodes. □

5.4. *Completeness result* Using the results of the previous subsections, we are able to show the completeness property.

CONSEQUENCE 5.2. *Let G be a CG over N and $\langle X,Y|Z\rangle \in T(N)$ such that $\neg\langle X,Y|Z\rangle_G^{sep}$. Then there exists a binary strictly positive G-Markovian distribution P over N such that $\neg\{\, X \perp\!\!\!\perp Y \,|\, Z\ [P]\,\}$.*

Proof. Whenever $\neg\langle X, Y|Z\rangle_G^{sep}$ then one can find $a \in X$ and $b \in Y$ such that $\neg\langle a, b|Z\rangle_G^{sep}$. Then one can choose a subgraph H described by Summary 5.1 and construct a binary probability distribution over N as described in Subsubsection 5.3.3. It is H-Markovian according to Lemma 5.6 for every choice of parameters. In case we choose nonzero parameters by Lemma 5.9 one has $\neg\{\, a \perp\!\!\!\perp b \,|\, Z\, [P]\,\}$ and hence $\neg\{\, X \perp\!\!\!\perp Y \,|\, Z\, [P]\,\}$ by the symmetry and decomposition properties of the independency model induced by P - see Subsection 2.1. Finally, the Markov property for H implies the Markov property for G according to Lemma 2.3. □

6. Appendix

6.1. *Outline of the subgraph choice procedure* In this appendix we show how a subgraph required in Summary 5.1 can be found. To give a global view we describe the main steps of the procedure performed in the following four subsections.

1. We introduce certain criteria by which to compare the 'complexity' of different active routes between two nodes and choose a route with minimal complexity in this sense - for details see Subsection 6.2. Moreover, we make a lot of observations about such a route.
2. We choose a free slide to every tail-terminal of a hit section of the chosen route.
3. A *regularization* of the chosen route will be made. That is such a modification that the new obtained route still has minimal complexity, and moreover it is a trail, its sections meet regularly, and the chosen slides are regular w.r.t. their host sections - for details see Subsection 6.4. Note that the regularization procedure is necessary to derive all the properties from Summary 5.1.
4. We choose a downroute for every free head-to-head section of the regularized route.
5. We verify that the regularized route, the chosen slides and the downroutes comply with all the requirements from Summary 5.1 and therefore the chosen subgraph composed of them solves our problem.

6.2. *Minimal complexity principle* Throughout the appendix we suppose that G is a CG over N and $\langle a, b|Z\rangle \in E(N)$ such that $\neg\langle a, b|Z\rangle_G^{sep}$ is given. Supposing that ρ is a route between a and b, which is active w.r.t. Z, the *section score* of ρ is the number of sections of ρ which are either hit by Z or head-to-head w.r.t. ρ. The *return score* of ρ is the number of pairs of distinct sections of ρ which meet each other. We say that ρ has *minimal complexity* if it has minimal section score within the class of active routes between a and b and moreover has minimal return score within the class of active routes between a and b with minimal section score.

Every pair of different sections α, β of the route ρ in G delimits a subroute of ρ consisting of sections between α and β including α and β. One can classify the

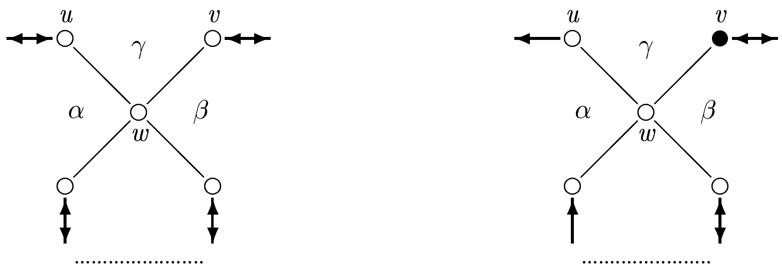

FIG. 15. *Illustrative pictures for the proofs of Observation 6.1 and Observation 6.2.*

terminals of α as follows: one of them is the *outward* terminal w.r.t. the subroute (that is, the endpoint of the subroute between α and β) and the other one is the *inward* terminal w.r.t. the subroute. Similarly the terminals of β can be classified. Note that this classification of terminals is independent of the classification according to the orientation of outgoing edges introduced in Subsection 4.1 (i.e. tail-terminals and head-terminals).

In all observations below we suppose that ρ is an active route between a and b which has minimal complexity. Note, that in proofs we will often utilize the following clear *transitivity principle*. If nodes u and v of G are connected by a free undirected route in G and there exists a free slide in G to u, then there exists a free slide in G to v.

Another *heritage principle* will be used when an old active route ρ in G will be replaced by a new route ϱ. If v is a node in N (usually a tail-terminal of a section of ϱ) which was a tail-terminal of a hit section of ρ, then the assumption that ρ is active implies that there exists a free slide in G to v.

OBSERVATION 6.1. *Different free sections of ρ do not meet each other.*

PROOF. Suppose for a contradiction that different free sections α and β share a node w. Let u (resp. v) be the outward terminal of α (resp. β) - see the left picture of Figure 15. Note that it may happen that $u = w$ or $v = w$. Then one can create from ρ a new route ϱ between a and b by replacing the subroute of ρ delimited by α and β by a new section γ made of the part of α between u and w and the part of β between w and v.

In case γ is a head-to-head section w.r.t. ϱ there exists a descending route copying ρ from w to a head-to-head section δ of ρ (possibly $\delta = \alpha$ or $\delta = \beta$). Thus, w has a descendant in Z and therefore ϱ is active w.r.t. Z and the section score of ϱ does not exceed the section score of ρ.

In case γ is not a head-to-head section it is not blocked since it is free. Thus, also in this case ϱ is active and the section score is again not raised by the

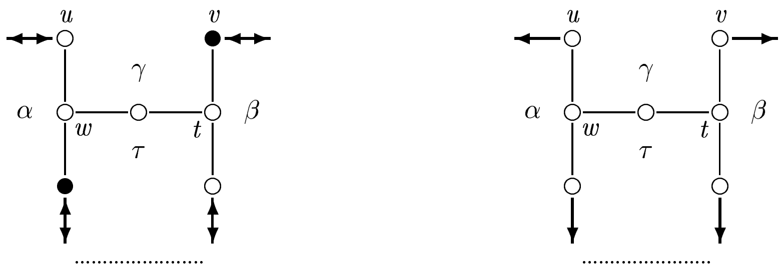

FIG. 16. *Illustrative pictures for the proofs of Observation 6.3 and Observation 6.4.*

described change.

However, it is evident that ϱ has lower return score than ρ, contradicting the assumption that ρ has minimal complexity. □

OBSERVATION 6.2. *A free head-to-tail section of ρ does not meet another section of ρ.*

PROOF. Suppose for a contradiction that α is a free head-to-tail section which shares a node w with a different section β and u (resp. v) are the outward terminals of α (resp. β) - see the right picture of Figure 15 (which depicts the subcase in which u is a tail-terminal). By the same procedure as in the proof of Observation 6.1 one can form a new route ϱ.

In case γ is a head-to-head section w.r.t. ϱ one can repeat the arguments from Observation 6.1 to show that ϱ is active and its section score does not exceed the section score of ρ.

In case γ is not a head-to-head section it suffices to show that γ is not blocked in ϱ. It is evident when γ is free. If γ is hit and u is its tail-terminal w.r.t. ϱ, then a free slide in G to u exists, made of nodes of α and the tail-terminal of the neighbouring section of α in ρ in direction to β (recall that α is a free head-to-tail section). In the case where v is a tail-terminal of γ the heritage principle ensures that there is a free slide in G to v since v is a tail-terminal of hit section β of ρ. Either, ϱ is active w.r.t. Z and the section score was not raised by the described change.

Finally, ϱ has lower return score than ρ which contradicts the assumption that ρ has minimal complexity. □

OBSERVATION 6.3. *There is no undirected route in G between nodes of different hit sections of ρ. In particular, different hit sections do not meet each other.*

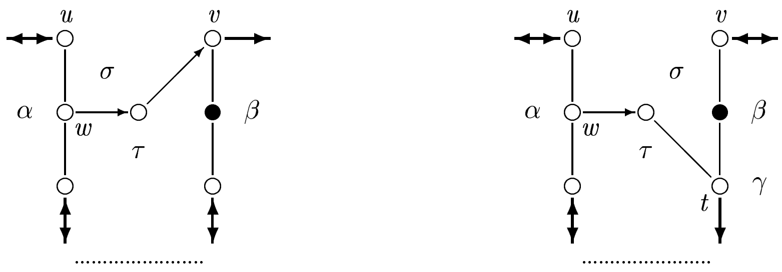

FIG. 17. *Illustrative pictures for the proof Observation 6.5.*

PROOF. Suppose for a contradiction that α and β are different hit sections with outward terminals u and v and that an undirected route τ connects a node w of α and a node t of β - see the left picture of Figure 16. Note that the case in which α and β meet each other occurs if $w = t$, i.e. τ consists of one node. One can create from ρ a new route ϱ by replacing its subroute between u and v by a new section γ made of the part of α between u and w, by τ, and by the part of β between t and v.

If γ is a head-to-head section w.r.t. ϱ, then it is not blocked by Z since it has evidently a descendant in Z.

If γ is not a head-to-head section, then it is either a free section or one can show that it is not blocked by the heritage principle. In either case, γ is not blocked.

Thus, ϱ is an active route with lower section score than ρ which contradicts the assumption. □

OBSERVATION 6.4. *There is no free undirected route in G between nodes of different free tail-to-tail sections of ρ.*

PROOF. Suppose for a contradiction that α and β are different free tail-to-tail sections with outward terminals u and v and that a free undirected route τ leads from a node w of α to a node t of β - see the right picture of Figure 16. One can create from ρ a new route ϱ by the same procedure as in the proof of Observation 6.3.

The new section γ is a free tail-to-tail section w.r.t. ϱ. Therefore ϱ is an active route with lower section score than ρ (there is at least one head-to-head section of ρ between α and β) which contradicts the assumption. □

OBSERVATION 6.5. *There is no free descending route in G from a node of a free section of ρ to a tail-terminal of a hit section of ρ.*

PROOF. Suppose for a contradiction that α is a free section and β is a hit section, u is the outward terminal of α, v the outward terminal of β, t is the inward terminal of β and that a free descending route τ leads from a node w of α to a tail-terminal of β. Thus, one can distinguish two basic cases: either τ leads to v or τ leads to t - see Figure 17. In both cases one can create from ρ a new route ϱ by replacing the subroute between u and v by a new subroute σ made of the part of α between u and w, by τ, and possibly by the whole section of β between t and v.

In case τ leads to v the whole subroute σ is composed of free head-to-tail or tail-to-tail sections. Therefore ϱ is active and has lower section score than ρ which contradicts the assumption.

In case τ leads to t the only section of σ which is not a free head-to-tail or tail-to-tail section is the last section γ, containing v. If γ is a head-to-head section w.r.t. ϱ, it is not blocked because it is hit. In the case where γ is not a head-to-head section we use the heritage principle to show that there exists a free slide in G to its possible tail-terminal v. The existence of a free slide in G to the other tail-terminal of γ (it may even be u!) can be shown by the transitivity principle: it is connected by an undirected free subroute of σ to t, and there exists a free slide to t (by the heritage principle). Thus, ϱ is active.

On the other hand, since t is a tail-terminal of β there exists a head-to-head section δ of ρ between α and β (otherwise the part of ρ from t to w together with τ forms a directed pseudocycle in G - see Lemma 2.2(ii)). The old subroute of ρ gave at least two points to the section score, while σ gives just one point. Hence, ϱ has lower section score than ρ what contradicts the assumption. $\square$

OBSERVATION 6.6. *There is no free descending route in G from a head-terminal of a hit section of ρ to a tail-terminal of a hit section of ρ.*

PROOF. Suppose for a contradiction that α is a hit section with a head-terminal and β a hit section with a tail-terminal, u (resp. v) is the outward terminal of α (resp. β), w (resp. t) is the possible inward terminal of α (resp. β) if it exists (note that in general one can have $\alpha = \beta$) and that a free descending route τ in G leads from a head-terminal of α to a tail-terminal of β. Thus, one can distinguish four basic cases according to the end-nodes of τ - see Figure 18. In all cases one can create from ρ a new route ϱ by replacing the subroute between u and v by a new subroute σ made possibly by α (if τ leads from w), by τ, and possibly by β (if τ leads to t).

In case τ leads from u to v (see the upper left picture of Figure 18, it may happen that $\alpha = \beta$) the new subroute σ is composed of free head-to-tail sections and therefore ϱ has lower section score than ρ which contradicts the assumption.

In case τ leads from u to t (see the upper right picture of Figure 18, here $\alpha \neq \beta$) only the section γ of σ containing v deserves attention. As it is hit, it is not blocked in the case in which it is a head-to-head section. Otherwise we use the heritage principle for v. The subroute σ adds just one point to the section score,

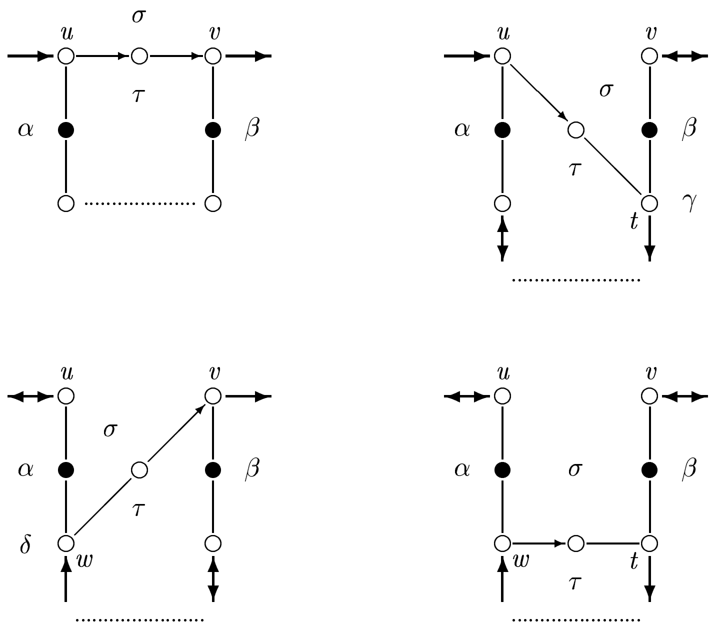

FIG. 18. *Illustrative pictures for the proof of Observation 6.6.*

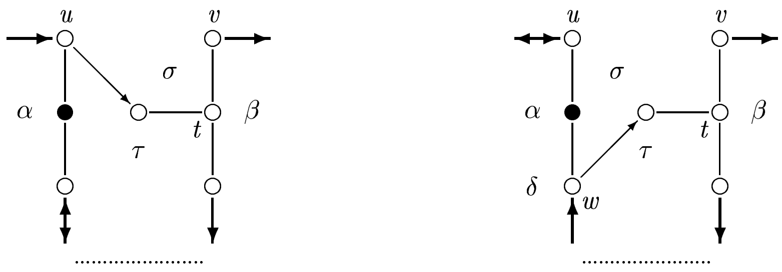

FIG. 19. *Illustrative pictures for the proof of Observation 6.7.*

while the the old subroute gave at least two points. Hence, the contradiction.

In case τ leads from w to v (see the bottom left picture of Figure 18, here $\alpha \neq \beta$) only the section δ containing u is interesting. To show that δ is not blocked one can use the heritage principle for u and the transitivity principle for the other tail-terminal (w has a free slide in G, simply the arrow of ρ heading to w). The decrease of the section score contradicts the assumption.

In the case where τ leads from w to t (see the bottom right picture of Figure 18) only the sections containing either u or v deserve attention. The arguments as to why they are not blocked can be repeated from the previous two cases. Moreover, there exists a head-to-head section in the subroute of ρ between α and β (otherwise the part of ρ from t to w can be prolonged by τ to a directed pseudocycle). Therefore the old subroute between u and v gave at least three points to the section score while σ gives at most two points, which contradicts the assumption. □

OBSERVATION 6.7. *There is no free descending route in G from a head-terminal of a hit section of ρ to a node of a free tail-to-tail section of ρ.*

PROOF. Suppose for a contradiction that α is a hit section with a head-terminal and β is a free tail-to-tail section, u (resp. v) is the outward terminal of α (resp. β), w the inward terminal of α and a free descending route τ leads from a head-terminal of α to a node t of β. Thus, one can distinguish two basic cases - see Figure 19. In both cases one can create from ρ a new route ϱ by a similar procedure as to that used in the proof of Observation 6.6.

In case τ leads from u (see the left picture) one can repeat the arguments used in the first case of the proof of Observation 6.6.

In case τ leads from w (see the right picture) one can verify that ϱ is active similarly to the third case in the proof of Observation 6.6. As there exists a head-to-head section in the subroute of ρ between α and β (otherwise there

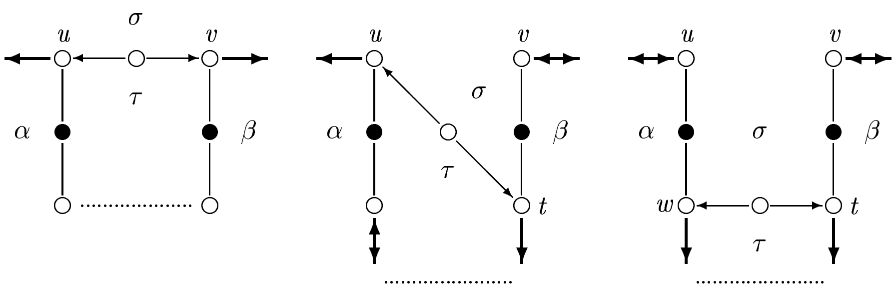

FIG. 20. *Illustrative pictures for the proof of Observation 6.8.*

exists a directed pseudocycle in G) the old subroute gave at least two points to the section score while σ gives just one point. Hence, the contradiction. □

OBSERVATION 6.8. *There is no free route without head-to-head sections in G between different tail-terminals of a hit section (or hit sections) of ρ.*

PROOF. Suppose for a contradiction that α and β (possibly $\alpha = \beta$) are hit sections and that a free route τ in G without head-to-head sections leads from a tail-terminal of α to a tail-terminal of β. Let u (resp. v) be the outward terminal of α (resp. β) (in the case where $\alpha = \beta$ simply u and v are the tail-terminals of the section). Denote by w (resp. t) possible inward terminals of α (resp. β) in the case where $\alpha \neq \beta$. One can distinguish three basic cases depending on whether outward or inward terminals are connected - see Figure 20. In all cases one can create from ρ a new route ϱ by a similar procedure to that used in the proof of Observation 6.6.

If τ connects u and v (see the left picture, here posssibly $\alpha = \beta$), then σ consists of free head-to-tail or tail-to-tail sections and no point is added to the section score by σ. But the old subroute gave at least one point to the section score, which contradicts the assumption.

If τ leads from u to t (see the middle picture, here $\alpha \neq \beta$), then only the section containing v deserves attention. The argument as to why σ is not blocked is similar the second case of the proof of Observation 6.5. The case when τ leads from w to v is covered by interchange of α and β.

If τ connects w and t (see the right picture), then only the sections of σ containing u and v are of interest. One can use the heritage and transitivity principles to show that they are not blocked and get the desired contradiction similarly to the fourth case of Observation 6.6. □

OBSERVATION 6.9. *There is no descending route in G which is free with the possible exception of its last node, and leads from a node of a free head-to-head*

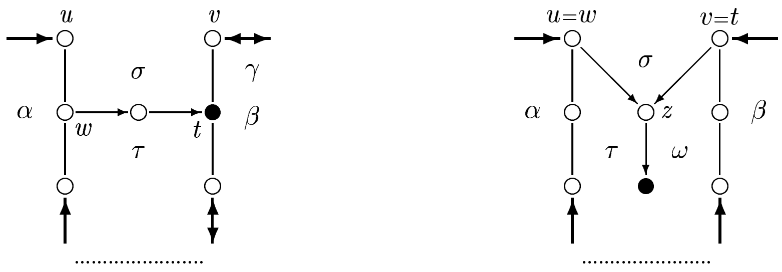

FIG. 21. *Illustrative pictures for the proofs of Observation 6.9 and Observation 6.10.*

section of ρ to a node of a different section of ρ. In particular, a free head-to-head section of ρ does not meet a different section of ρ.

PROOF. Suppose for a contradiction that α is a free head-to-head section, β a different section, u (resp. v) the outward terminal of α (resp. β), a descending route τ in G leads from a node w of α and a node t of β, and all nodes of τ with the possible exception of t are outside Z - see the left picture of Figure 21. Note that the case when α meets β corresponds to the situation $w = t$. Then one can create from ρ a new route ϱ by replacing the subroute between u and v by a new subroute σ made possibly of the part of α between u and w (if $u \neq w$), by τ, and possibly by the part of β between t and v (if $t \neq v$). With the possible exception of the section γ containing v all sections of σ are free head-to-tail sections.

If γ is a head-to-head section w.r.t. ϱ, then either β is a head-to-head section w.r.t. ρ or there exists a head-to-head section δ of ρ between α and β (otherwise the part of ρ from t to w together with τ forms a directed pseudocycle). In either case γ is active. Therefore the old subroute between u and v gave at least two points to the section score while σ adds at most one point, which contradicts the assumption.

If γ is not a head-to-head section, then either it is free (and the section score decreases) or one can use the heritage principle to show that it is not blocked. Then σ adds just one point to the section score while the old subroute gave at least two points. In either case, it contradicts the assumption. □

OBSERVATION 6.10. *Descending routes in G from nodes of different free head-to-head sections of ρ to nodes in Z, which are free with the exception of their last nodes, do not meet each other.*

PROOF. Suppose for a contradiction that α and β are different free head-to-head sections, u (resp. v) is the outward terminal of α (resp. β), τ is a descending route from a node w of α to a node in Z, ω is a descending route from a node t of β to a node in Z, and τ and ω share a node z - see the right picture of Figure 21.

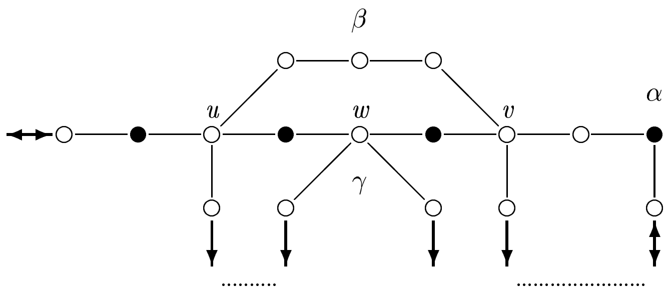

FIG. 22. *An illustrative picture for the proof of Observation 6.11.*

Then one can form from ρ a new route ϱ by replacing the subroute of ρ between u and v by a new subroute σ consisting of the part of α from u to w (if $u \neq w$), the part of τ from w to z, the part of ω from z to t and the part of β from t to v (if $t \neq v$).

All sections of σ with exception of the section containing z are free head-to-tail sections. The section which contains z is a head-to-head section w.r.t. ϱ and has a descendant in Z. Therefore σ adds just one point to the section score while the old subroute gave at least two points. That contradicts the assumption. □

OBSERVATION 6.11. *All the occurrences of nodes of a free tail-to-tail section of ρ within a hit section α of ρ cannot be confined in the subroute of α between the occurrences of nodes of a different tail-to-tail section within α.*

PROOF. Suppose for a contradiction that α is a hit section, β and γ are different free tail-to-tail sections, u and v is the first and last occurence of a node of β within α, at least one node w occurs both in γ and α, and all the occurences of nodes of γ in α are between u and v - see Figure 22. Then one can form from ρ a new route ϱ by replacing the subroute of ρ within α between u and v by a new subroute σ made of the part of β from u to v.

The revised section α is not blocked in ϱ for the same reason that the original section α was not blocked in ρ. Moreover, the section score of ϱ does not exceed the section score of ρ. However, as β does not meet a section other than α (see Consequence 6.5 below which is proved without this Observation 6.11) ϱ has lower return score than ρ (γ and α do not meet each other after the change) which contradicts the assumption that ρ has minimal complexity. □

6.3. *The choice of the route and free slides* Given a CG G over N and $\langle a,b|Z\rangle \in E(N)$ with $\neg\langle a,b|Z\rangle_G^{sep}$ our first construction step is to choose a route between a and b which is active w.r.t. Z and has minimal complexity in the sense described in Subsection 6.2. Let us denote it by ρ. Although it could be modified within sections in the following Subsection 6.4 its arrows between sections will remain fixed. The first basic property of the chosen route is formulated here.

CONSEQUENCE 6.1. *If α, β are different sections of a minimal complexity active route which share a node, then up to the mutual interchange of α and β the following conditions hold.*

(i) *α is a free tail-to-tail section,*
(ii) *β is a hit section,*
(iii) *the shared nodes do not belong to any terminal free zone of β.*

PROOF. Observation 6.1 excludes the possibility that both α and β are free, while Observation 6.3 excludes the possibility that both α and β are hit. Let's suppose concretely that α is free and β hit. By Observation 6.2 α is not a head-to-tail section, by Observation 6.9 α is not a head-to-head section. So, α is a tail-to-tail section. Finally, by Observation 6.5 α cannot meet a tail-terminal free zone of β and by Observation 6.7 α cannot meet a head-terminal free zone of β. □

The second construction step is to choose for every tail-terminal of a hit section of ρ a free slide in G to that tail-terminal. To avoid ambiguity of the procedure the following property is useful.

CONSEQUENCE 6.2. *Terminal free zones for different tail-terminals of hit sections (resp. of a hit section) of a minimal complexity active route do not meet each other. In particular, no node of G can be twice (or more times) a tail-terminal of a hit section of ρ.*

PROOF. This follows directly from Observation 6.8. □

The chosen route ρ is active w.r.t. Z by the assumption, and therefore there exists at least one free slide to every tail-terminal of a hit section of ρ. We choose and fix just one free slide for every such tail-terminal. Owing to Consequence 6.2 'tail-terminal' zones in ρ are disjoint and therefore no slide in G can lead to distinct 'tail-terminal' zones. Thus, every (free) slide in G has uniquely determined its host section in a minimal complexity active route ρ (more exactly, the 'tail-terminal' free zone of the host section is uniquely determined). And, of course, one cannot ascribe the same free slide to different tail-terminals. Moreover, one can easily derive the following stronger conclusion.

CONSEQUENCE 6.3. *The (chosen) free slides to different tail-terminals of hit sections (resp. of a hit section) of a minimal complexity route do not meet each other. Any chosen free slide does not meet a free section of the route or a terminal free zone of a hit section of the route (except the corresponding terminal free zone of the host section).*

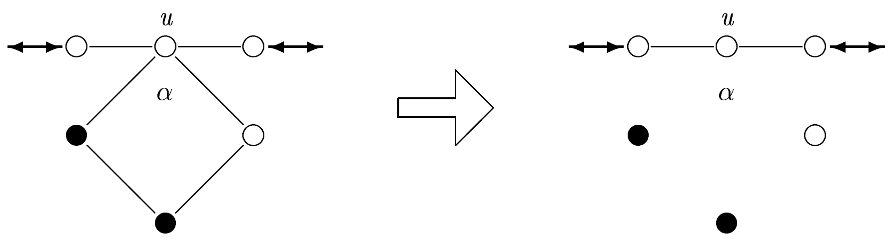

FIG. 23. *Shortening of sections.*

PROOF. The first claim follows from Observation 6.8, the second one from Observation 6.5 and the third one from Observations 6.6 and 6.8 (depending on whether one considers a head-terminal free zone or a tail-terminal free zone). □

6.4. *Regularization* The third constuction step is regularization of ρ, that is, a series of possible changes within sections. The changes will be performed in the following three stages.

- Sections of ρ will be possibly shortened to consist of distinct nodes.
- Hit sections with tail-terminals will be possibly modified to have the chosen free slides regular with respect to their host sections.
- Hit sections which meet tail-to-tail free sections will be possibly modified to meet regularly those free tail-to-tail sections.

6.4.1. *Shortening of sections* Supposing α is a section of ρ and u a node of α which is repeated in α we omit the subroute of ρ (within α) between the first and the last occurence of u in α and replace it by a single occurence of u - see Figure 23. Evidently, the shortened section α has the same terminals as the original section α, and therefore the modified route will also be active and the section score will not be raised. As ρ was supposed to have minimal section score, the section score remains unchanged. Similarly, the return score is saved by the described change and the modified route also has minimal complexity. Nevertheless, the result of all such modifications is a route having sections consisting of distinct nodes.

6.4.2. *Regularization of hit sections with respect to free slides* Supposing u is a tail-terminal of a hit section of ρ (already with sections consisting of distinct nodes), α is the hit section having u as its tail-terminal (uniquely determined by Consequence 6.2) and ε the free slide chosen for u in Subsection 6.3. Then we find the last node w of α in the direction away from u which belongs to ε. Surely, w is

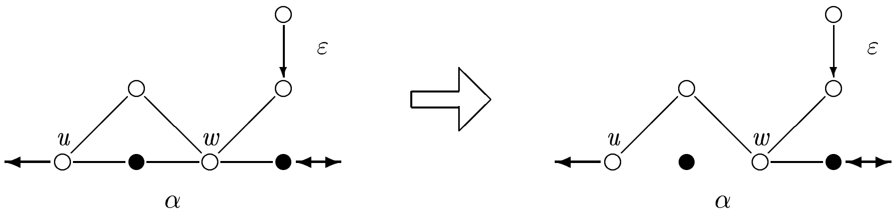

FIG. 24. *Regularization with respect to free slides.*

not the top node of ε (otherwise a directed pseudocycle is formed by ε and by a part of α). Thus, one can replace the subroute of ρ between u and w within α by the corresponding part of ε - see Figure 24 (that is under condition that ε is not already regular with respect to α). As only α could possibly be changed, but not its terminals, the modified route will also be active. The section score is surely not raised by this modification. Owing to the assumption that ρ has minimal complexity the section score keeps the same value. To see that the return score was not raised by the described change one needs the following property.

CONSEQUENCE 6.4. *There is no free undirected route in G between a tail-terminal of a hit section of a minimal complexity route and a different section of the route.*

PROOF. In the case that the other section is hit it follows from Observation 6.3, in the case that it is free from Observation 6.5 (note that every undirected route is a descending route). □

Indeed, if a section β of the modified route (different from α) meets α in the modified route, then β is a section of the original route ρ and by Consequence 6.4 it cannot meet ε in a non-top node. Hence β also has to meet α in the original route ρ. Thus, the return score was not raised, so by the assumption about ρ it is saved and the modified route also has minimal complexity.

As w is the last occurence of a node of ε in α the rest of α after w does not meet the new part of α. Moreover, both α in the original route ρ and ε consist of distinct nodes. Therefore also the modified section α will consist of distinct nodes. Owing to Consequence 6.3 the node w can never occur in a part of α formed from a (free) slide to another tail-terminal of α during a prior regularization step. Thus, all formerly regularized slides remain regular and the result of all these modifications will be a route with sections consisting of distinct nodes such that all chosen slides are regular with respect to their host sections.

6.4.3. *Regularization of hit sections with respect to free tail-to-tail sections* To ensure feasibility of the procedure below the following property is needed.

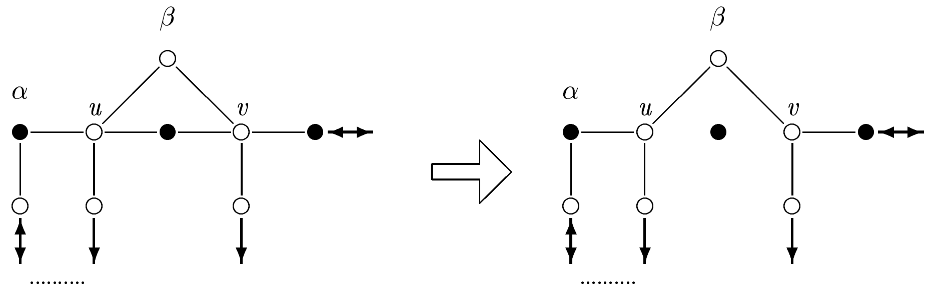

FIG. 25. *Regularization with respect to free tail-to-tail sections.*

CONSEQUENCE 6.5. *A free tail-to-tail section of a minimal complexity route meets at most one other section of the route.*

PROOF. Suppose that a free tail-to-tail section α meets two different sections β and γ. By Observation 6.1 both β and γ is hit. But then an undirected route in G between nodes of β and γ composed of nodes of α exists which contradicts Observation 6.3. □

Thus, supposing α is a free tail-to-tail section of ρ which meets another section of ρ, the (hit) section β which meets α is uniquely determined. We find the first and the last occurences of a node of α in β and denote them by u and v and possibly replace the subroute of ρ between u and v within β by the corresponding part of α - see Figure 25 (that is under condition that α and β do not meet regularly already). As only β can possibly be affected by the change, but not in its terminals, the modified route will remain active. Moreover, the section score will not be raised and therefore, by the assumption about ρ, the value of the section score is saved. A similar conclusion holds for the return score. If a section γ of the modified route (different from β) meets β in the modified route, then it is a section of the original route ρ and by Consequence 6.5 cannot share a node with α unless it is α. Hence γ must meet β also in the original route ρ and the modified route still has minimal complexity.

Owing to the definition, the new part of β does not meet the remaining original parts of β. As sections were already composed of distinct nodes, the modified section β will also consist of distinct nodes.

Consequence 6.3 implies that the nodes u, v (belonging to a free section) can never occur in the terminal parts of β which were possibly made of nodes of (chosen) free slides during a former regularization procedure with respect to the slides. Thus, the chosen slides will also remain regular with respect to their host sections after the described change.

As one hit section β can meet several free tail-to-tail sections we have also to verify that the described change saves the results of a former regularization procedure with respect to another free tail-to-tail section which meets β. Owing to Observation 6.1 no node of α can belong to another free tail-to-tail section, in particular neither u nor v can be inside the part of β which was possibly formed from the nodes another free tail-to-tail section during a prior regularization procedure. By Observation 6.11 a formerly regularized part of β also cannot occur between u and v in β. Thus, the interval of the original section β between u and v is does not meet a formerly regularized region of β. Hence, the results of all former regularizations will be saved.

Thus, after all described changes of the modified route one has an active route with minimal complexity, its sections consist of distinct nodes, the chosen slides are regular with respect to their host sections and its sections meet regularly (recall Consequence 6.1). To see that the regularized route is a trail we realize that its arrows cannot be repeated as otherwise two different sections share a tail-terminal which contradicts Consequence 6.1.

6.5. *The choice of downroutes and verification of conditions from Summary 5.1* There exists at least one downroute for every free head-to-head section of the chosen active trail. Our fourth construction step is to choose and fix just one downroute for every free head-to-head section of the (regularized) trail. As different free head-to-head sections do not share a node by Observation 6.1 one cannot ascribe the same downroute to different free head-to-head sections. Moreover, one can easily derive the following properties.

CONSEQUENCE 6.6. *Downroutes chosen for different free head-to-head sections do not meet each other. Every chosen downroute does not meet any chosen slide or a section of a minimal complexity route except its host section.*

PROOF. The first claim follows from Observation 6.10, the second claim can be derived from Observation 6.9. □

The fifth construction step is to verify that the regularized route, the chosen slides and the chosen downroutes satisfy all conditions from Summary 5.1.

Condition (a) follows from Consequence 6.6.

As concerns condition (b) the regularity of the chosen free slides was ensured in Subsection 6.4, by Consequences 6.3 and 6.6 the chosen slides can only meet hit sections outside their terminal free zones (with the exception of the host section). If the slide meets a section in a non-top node, then by Observation 6.3 it is the host section. Otherwise they share just the top node, necessarily in a middle free zone, and the section is uniquely determined by Observation 6.3 (otherwise two different hit sections of ρ share a node).

As concerns condition (c), the regularity was ensured in Subsection 6.4, the rest can be derived from Consequences 6.1 and 6.5.

As concerns condition (d), the slide uniqueness follows from Observation 6.8, the section uniqueness from Observation 6.4 and the combined uniqueness from Observation 6.5.

Condition (e) is evident.

Condition (f) can be verified by an analysis of possible connectivity components in H using Consequences 6.1, 6.3 and 6.6:

- If a node does not belong to a downroute, a slide or a section, then it is isolated in H and therefore its connectivity component is a singleton without parents.
- If a node belongs to a downroute, then either its connectivity component in H is a simple path with one parent or it belongs to the connectivity component of the host section of the downroute.
- The top node of a slide either forms a singleton connectivity component without parents in H or it belongs to a connectivity component of another hit section.
- A non-top node of a slide belongs to the connectivity component of the host section.
- If a node belongs to a free head-to-tail section, then its connectivity component in H consists of a simple path with one parent.
- If a node belongs to a free tail-to-tail section, then its connectivity component in H is a simple path without parents or it belongs to the connectivity component of a hit section.
- If a node belongs to a free head-to-head section, then its connectivity component is a simple path between the terminals of the section with a possible branch made of nodes of the corresponding downroute.
- If a node belongs to a hit section, then its connectivity component consists of a simple path between its terminals and possible branches made of slides or parts of free tail-to-tail sections which meet the hit section (note that no cycle can appear as the route ρ was regularized!).

As concerns condition (g), the preceding analysis says that two arrows may be directed only to a connectivity component of H which corresponds to a hit or head-to-head section α. An entering arrow can either belong to a free slide for a tail-terminal of α (then the parent node of the connectivity component is the top node of the slide) or it can belong to the route (then the parent node is a tail-terminal of a neighbouring section of α). Thus, the parent nodes differ by Consequences 6.1 and 6.3. They cannot be joined by an edge of a chosen slide or a downroute by Consequences 6.3 and 6.6. Supposing that the parent nodes are joined by a line of a section ξ one can derive by Consequence 6.3 (the case of the top node of a slide) resp. by Consequences 6.1 and 6.5 (the case of tail-terminals) that the section ξ is hit and the considered line belongs to one of its middle free zones. But that situation contradicts to condition (d) which has already been verified. Suppose the parent nodes are joined by an arrow $u \to v$ of ρ (say, originating from a section β) Consequence 6.3 excludes u belonging to a chosen free slide. But then at least one of the neighbouring sections of α, say γ, meets the section β at u. The case $\beta \neq \gamma$ means that they meet in their terminals which contradicts Consequence 6.1. The nasty possibility that $\beta = \gamma$ can be excluded as follows. Of course, the arrow $u \to v$ cannot coincide with the arrow from β to α as otherwise v is simultaneously a parent of the component and belongs to the component (resulting in a directed pseudocycle). Thus, the only remaining possibility is that β $(= \gamma)$ is a singleton tail-to-tail section. Then v is a head-terminal of a further section δ (in direction from α to β). Thus, v cannot belong to a free slide by Consequence 6.3. Hence, v has to be a tail-terminal of the other neighbouring section of α, say of η (i.e. different from β). Since $\eta \neq \delta$ Consequence 6.1 leads to a contradiction. Thus, the parent nodes cannot be joined by an edge in H.

To verify (h) suppose for a contradiction that $\lambda_1 : u \to t_1 - \ldots - t_s \leftarrow v$, $s \geq 1$ and $\lambda_2 : u \to w_1 - \ldots - w_r \leftarrow v$, $r \geq 1$ are complexes in H having different nodes u, v as parents. The region of any complex in H must belong to a connectivity component of H with two parents. Supposing that $\{t_1, \ldots, t_s\}$ and $\{w_1, \ldots, w_r\}$ belong to different connectivity components of H, which correspond to different sections β and γ of ρ by Consequence 6.3 none of the arrows $u \to t_1$ and $u \to w_1$ can belong to a chosen free slide. Thus, the arrows belong to ρ and u should be twice a tail-terminal of a section, namely of neighbouring sections of β and γ. Therefore, by Consequence 6.1 the neighbouring sections of β and γ containing u must coincide. The only possibility is that the the singleton $\{u\}$ forms a free tail-to-tail section between β and γ. However the same conclusion can be derived for v which implies the contradictory conclusion $u = v$. Thus, the regions of λ_1 and λ_2 must belong to the same connectivity component C of H. Necessarily $t_1 = w_1$ as otherwise different arrows $u \to t_1$, $u \to w_1$ and $v \to t_s$ enter C which contradicts the condition (g). Similarly $t_s = w_r$. However according to (f) and Lemma 2.1(ii) there exist just one undirected path between $t_1 = w_1$ and $t_s = w_r$ in H. Hence $\lambda_1 = \lambda_2$.

6.6. *Proof of Lemma 5.1* Throughout this subsection we suppose that H is the chosen subgraph for $\langle a, b|Z\rangle \in E(N)$ mentioned in Summary 5.1 and $u_1, \ldots, u_n$, $n \geq 3$ is the corresponding sequence of segment delimiters introduced in Subsubsection 5.2.3. Our proof is based on an analysis of special paths in H^{mor}. By a *basic path* in H^{mor} (resp. in H) we will mean any path in H^{mor} (resp. in H) between nodes of $u_1 \ldots u_n Z$ whose inner nodes are outside $u_1 \ldots u_n Z$.

The first claim of Lemma 5.1 says that for $i = 1, \ldots, n-2$ every path in H^{mor} from $u_1 \ldots u_i$ to $u_{i+2} \ldots u_n$ is hit by $u_{i+1}Z$. To reach this conclusion it suffices to verify that every path in H^{mor} from u_i to $u_{i+2} \ldots u_n$ is hit $u_{i+1}Z$ (for each $i = 1, \ldots, n-2$). One can show this by verifying that in every case which can occur for u_i every basic path from u_i in H^{mor} leads either to a node in Z, or to u_{i+1}, or to u_{i-1} (when $i \geq 2$). We will give a complete justification of our arguments on the basis of Summary 5.1 only in several initial cases, just to show how a detailed proof can be carried out. The justification of analogous arguments in later cases would be nothing but a mechanical repetition of considerations already mentioned from in the initial cases and will be omitted. Several final cases will be left to the reader completely.

6.6.1. *Basic paths from the first node* First, we show that every basic path in H^{mor} from u_1 leads to u_2Z. One can distinguish several possibilities for the first section α of the considered trail which has u_1 as its tail-terminal.

If α is a free head-to-tail section, then no other edge in H enters a node of α except lines of α and the arrow from the tail-terminal u_2 of the next section to the head-terminal of α. Indeed, such an edge belongs either to the trail τ, or to a free slide for some tail-terminal of a hit section of τ, or to a downroute for a free head-to-head section of τ. It does not belong to a downroute as otherwise the downroute meets α which contradicts Summary 5.1(a). It does not belong to a slide as otherwise the slide meets α which contradicts Summary 5.1(b). If the edge belongs to τ (and it is not of the type described above), then α meets another section of τ which contradicts Summary 5.1(c). Moreover, no virt in H^{mor} can enter a node of α. Indeed, if a virt in H^{mor} enters a node $t \in N$, then there exists an arrow in H outgoing from t and this is not the case. Thus, every basic path from u_1 in H^{mor} leads to u_2 in this case.

If α is a free tail-to-tail section with at least two nodes, then Summary 5.1(a)-(c) implies that the only edges in H (except the lines of α and the arrow of τ outgoing from the other tail-terminal u_2 of α) which can possibly enter α, are lines of a middle free zone of a hit section of τ (uniquely determined in that case). But then no further edge in H enters a node of the middle free zone by Summary 5.1(d). Thus, every basic path in H from u_1 leads either to a node in Z (a limit of the middle free zone) or to the other tail-terminal u_2 of α. With the exception of u_2 the nodes of such a path cannot be entered by a virt in H^{mor}. Hence, the claim can be extended to basic paths in H^{mor}.

If α is a singleton free tail-to-tail section, then the arrow from u_1 enters the next section β and one can distinguish four subcases depending on the type of β. However, in each subcase the arguments concerning a possible middle free zone which meets α can be repeated.

- If β is a free head-to-tail section, then by Summary 5.1(a)-(c) no other edge in H enters β (except those implicitly considered). In particular, the connectivity component of H containing the nodes of β has just one parent u_1. This implies that no virt in H^{mor} can enter u_1 and every basic path in H^{mor} from u_1 leads either to Z or to the tail-terminal u_2 of β.
- If β is a hit head-to-tail section, then the arrow from u_1 to β enters either a node in Z or a head-terminal free zone of β. In the latter case one can show that no other edge in H^{mor} enters the head-terminal free zone (except those implicitly considered). However, there exists a virt in H^{mor} between u_1 and the top node w of the slide to the tail-terminal of β. The slide then leads to its heel u_2 and the only possible branches in H^{mor} occur when a middle free zone of another hit section contains w. In either case, every basic path in H^{mor} from u_1 leads to Z or u_2.
- If β is a hit head-to-head section, then one can repeat the arguments from

previous case with the modification that the virt from u_1 leads directly to the tail-terminal u_2 of the section following after β.

- If β is a free head-to-head section, then one can use the same procedure, but instead of a head-terminal free zone of β one considers the whole free section β and a downroute for β to a node in Z. By standard arguments one can show that no other edge in H^{mor} can enter β or the downroute (except those implicitly considered).

If α is a hit section and the heel u_2 of the slide to u_1 differs from u_1, then no other edge in H^{mor} enters the path between u_1 and u_2 (except those implicitly considered).

If α is a hit tail-to-tail section and u_1 is also the heel of the slide for u_1, then the only two edges in H^{mor} entering u_1 are a line of α and an edge of the slide. The line of α leads in H^{mor} to a node in Z while the slide leads to its top node w. The only other edges of H_{mor} possibly entering w are the lines of a middle free zone of another hit section which lead to a node in Z, and the virt to the top node t of the slide to the other tail-terminal of α. With the proviso of a possible middle free zone of a hit section which can meet t the other slide leads to its heel u_2.

If α is a hit head-to-tail section and u_1 the heel of the corresponding slide, then one can repeat the procedure from the previous case with the modification that the virt from w leads directly to the tail-terminal u_2 of the next section β.

6.6.2. *Basic paths from the other nodes* The verification of the fact that every basic path from u_i $(2 \leq i \leq n-2)$ leads to $u_{i-1}u_{i+1}Z$ is analogous and will be left to the reader. The main difference is the larger number of cases as one has to distinguish several possible 'directions' out of u_i (up to 6 edges in H^{mor} may enter u_i).

If u_i is a tail-terminal of a section α and an outgoing arrow from u_i enters a section β (either the next one or the preceding one), then one can distinguish four cases for β depending whether it is free or hit, and a head-to-tail or a head-to-head section. This is completely analogous to the 4 subcases of the situation from the preceding subsubsection, where u_1 forms a singleton free tail-to-tail section. The analysis used there can also be used here to show that basic paths in H^{mor} from u_i which start either with the outgoing arrow or with the corresponding (possible) virt lead to $u_{i-1}u_{i+1}Z$.

If u_i is a tail-terminal of a section α and one considers basic paths in H^{mor} from u_i which start with another edge in H^{mor}, then one can distinguish (up to) nine cases analogous to all nine cases of the situation from the preceding paragraph. The arguments of that analysis can repeated here, too.

If u_i is a heel of a slide for a different tail-terminal (either u_{i-1} or u_{i+1}) of a hit section α, then one can distinguish 3 cases: either to consider the basic path from u_i which starts with the line of α in the direction towards the tail-terminal, or to consider the other directions (which involve also the path copying the slide in direction towards its top node) and to distinguish two subcases depending on whether α is a head-to-tail or a tail-to-tail section.

6.6.3. *The second claim of Lemma 5.1* To derive the second claim of the lemma from the first one it suffices to realize that for every $i = 1, \ldots, n-2$ the graph H_T^{mor} where $T = \mathrm{an}_H(u_i u_{i+1} u_n Z)$ is a subgraph of H^{mor} and therefore every path from u_i to u_n in H_T^{mor} is also a path from $u_1 \ldots u_i$ to $u_{i+1} \ldots u_n$ in H^{mor}.

7. Conclusions

7.1. *Main results* We can summarize the obtained results as follows.

THEOREM 7.1. (equivalence of criteria)
Let G be a CG over N, $\mathcal{B}$ be a chain for G, and $t \in T(N)$. Then, the following conditions are equivalent:

(i) t is represented in G according to the moralization criterion,
(ii) t is represented in G according to the separation criterion,
(iii) t belongs to the graphoid closure of the input list associated with G and $\mathcal{B}$.
(iv) t belongs to the probabilistic closure of the input list associated with G and $\mathcal{B}$ w.r.t. the class of strictly positive (discrete) probability distributions.

PROOF. Consequence 4.1 says $(i) \Leftrightarrow (ii)$, Consequence 3.1 $(i) \Leftrightarrow (iii)$, the implication $(iii) \Rightarrow (iv)$ follows from basic properties of the graphoid closure mentioned in Subsection 2.1. The implication $(iv) \Rightarrow (ii)$ can be proved by contradiction: if $\neg\langle X, Y|Z\rangle_G^{sep}$ then by Consequence 5.2 there exists a strictly positive G-Markovian probability distribution P with $\neg\{\, X \perp\!\!\!\perp Y \mid Z \; [P] \,\}$. By Consequence 3.2 P satisfies the triplets from the input list. Thus, $\langle X, Y|Z\rangle$ does not belong to the probabilistic closure of the input list w.r.t. the considered class of distributions, which contradicts (iv). □

The previous theorem solves the open question of completeness of the moralization criterion, formulated in [17] and [9], with a positive answer. The equivalence $(iii) \Leftrightarrow (iv)$ of the previous theorem can be interpreted also as follows: graphoid properties have been shown to be complete for input lists w.r.t. the class of strictly positive probability distributions. Moreover, $(i) \Leftrightarrow (iii)$ and $(i) \Leftrightarrow (iv)$ imply that both the graphoid closure and the mentioned probabilistic closure of the input list does not depend on the particular choice of chain.

However, to justify the use of CGs as models of conditional independence structures for strictly positive discrete probability distributions one needs to prove the following stronger claim which generalizes analogous results for DAGs and UGs from [11, 12].

THEOREM 7.2. (strong completeness)
If G is a CG over N, then there exists a strictly positive probability distribution over N which satisfies exactly those triplets in $T(N)$, which are represented in G

according to one of the above mentioned criteria. In particular, every CG-model is a probabilistic independency model.

PROOF. The key property of probabilistic independency models is the following one. Whenever P_1 and P_2 are two strictly positive probability distributions over N (say P_1 defined on $\prod_{i\in N} \mathbf{X}_i$ and P_2 defined on $\prod_{i\in N} \mathbf{Y}_i$) then there exists a strictly positive probability distribution $P_{1,2}$ over N such that $X \perp\!\!\!\perp Y \,|\, Z\ [P_{1,2}]$ iff $\{\ X \perp\!\!\!\perp Y \,|\, Z\ [P_1]\ \&\ X \perp\!\!\!\perp Y \,|\, Z\ [P_2]\ \}$. For example, one can define $P_{1,2}$ on $\prod_{i\in N} \mathbf{X}_i \times \mathbf{Y}_i$ by the formula

$$P_{1,2}([\mathbf{x}_i, \mathbf{y}_i]) \;=\; P_1([\mathbf{x}_i]) \cdot P_2([\mathbf{y}_i]) \qquad \text{for } [\mathbf{x}_i, \mathbf{y}_i] \in \prod_{i\in N} \mathbf{X}_i \times \mathbf{Y}_i.$$

Thus, if one considers the list of all triplets $t_1, \ldots, t_n \in T(N)$ which are not represented in G, by Theorem 7.1 for each t_i there exists a strictly positive probability distribution P_i over N satisfying all triplets represented in the graph, but not t_i! By repeated application of the construction above one can obtain the distribution $P_{1,\ldots,n}$ which has the desired property. $\square$

REMARK. Note that both the concept of strong completeness and the construction used in the preceding proof is taken from [12]. The reader should be aware of the fact that the desired distribution obtained may have a pretty large domain $\prod_{i\in N} \mathbf{X}_i$. The construction presented does not help in the case where one needs to construct such a distribution on a domain $\prod_{i\in N} \mathbf{X}_i$ prescribed apriori. Let us note in this context that in the framework of DAGs a (non-constructive) proof of the existence of a distribution given a prescribed DAG-model and a domain was given in [21].

7.2. *Discussion* Let us start the discussion with a (slightly modified) question raised by a reviewer.

Do CGs help us to build better models?

Well, we do not feel entitled to answer the question on such a level of generality. We hope that future development in statistics and AI will confirm the usefulness of CG-models. Perhaps we can give here a few arguments in favour of CG-models.

The first argument is that CGs make it possible to describe a wider class of probabilistic independency models involving both UG-models, DAG-models, and recursive causal models from [13]. And not only these models. They are CG-models which cannot be descibed by means of the above mentioned graphs. Thus CGs enlarge the applicability of graphical models for description of conditional independence structures of discrete probability distributions. In particular, they allow a more precise inner approximation of general probabilistic models (graphoids) than the classic graphs (UGs and DAGs).

The second argument is that in the area of probabilistic expert systems CGs may perhaps to help represent more precisely structural information obtained from experts. In fact this is one of possible ways of interpretation of CG-models based on the above mentioned factorization property. We omit the details of that interpretation [27] and explain just the main issue. Suppose that we have a group of experts, and each of them has his/her own area of competence (consisting of a set of variables). The areas of competence are disjoint and cover all the area of (our) interest. The experts are asked to describe structural relationships among variables which they think are relevant to the variables within their areas of competence. Thus, each of them is supposed to provide an UG having as the set of nodes both the area of competence but possibly an area of influence (consisting of those variables outside the area of competence which the expert indicates as relevant to some variables within his/her area of competence). To prevent discrepancy between different experts we order their areas of competence and accept the rule that experts' areas of influence will be restricted to the variables taken from the preceding areas of competence (because the relationships to the variables taken from the 'subsequent' areas of competence will be described by 'subsequent' experts). Our task is to put together pieces of structural information from the experts and represent it properly. Then a certain CG, constructed on the basis of UGs provided by the experts, provides an exact natural description of the overall amount of structural information.

The third argument is that CGs provide not only a unification of classic graphical models but also (perhaps surprisingly) exhibit more elegant mathematical properties (we have in mind the original interpretation of CGs given in [16]). Frydenberg [9] showed that every class of Markov equivalent CGs contains a distinguished representative which he called the *largest chain graph* (it has the greatest number of lines within the class of Markov equivalent CGs). Note that this result has no analogy in the case of DAGs: Markov equivalent DAGs have no distinguished members and have to be represented by more general mathematical objects, like for example, essential graphs [3]. This fact causes many complications for search procedures within the class of DAG-models (that is, the procedures for estimation of the most suitable DAG-model approximation on the basis of statistical data). However, in the case of CGs one can simply represent every CG-model by the corresponding largest CG and search within the class of the largest CGs only. One just needs a suitable graphical characterization of the largest CGs. Note that the paper [26] describes a recovery algorithm which on the basis of a given CG-model finds the corresponding largest CG. Moreover, an algorithm which converts every CG into the corresponding largest CG is presented there and this leads to a certain graphical characterization of the largest CGs. However, an even more natural and simpler graphical characterization of the largest CGs was found very recently [29]. Thus, it may appear to be a more convenient approach to represent even the classes of Markov equivalent DAGs by the corresponding largest CGs!

7.3. *Open questions* Let us comment on the results achieved in this paper and formulate the corresponding open problems. The new separation criterion, based on the concept *c-separation* has its own significance. For example, it easily

implies that every CG-model satisfies the following property from [22]
$[\langle X,Y|Z\rangle \ \& \ \langle X,W|Z\rangle] \quad \longrightarrow \quad \langle X,YW|Z\rangle$ composition,
which may be complicated to verify using the moralization criterion. One of the reviewer's asked whether c-separation could be interpreted by simple stories as in case of UGs and DAGs. In fact, we have not tried to elaborate such a story so far although we have a few vague ideas as to how (possibly) to develop such interpretation. Of course, such a story would be more complicated than in case of DAGs because c-separation is a more complicated concept than d-separation (no wonder, that's a natural price for more general models). Nevertheless, we admit that such a story is important especially for students and consider this question as one of the open problems.

Another open question concerns the computational complexity of the c-separation criterion. Perhaps it can be effectively used as the basis of an algorithm for testing independency statements represented in a CG provided that one finds a clever method of implementation. On the basis of the alternative formulation of the separation mentioned in the end of Section 4 one can construct a locally performed propagation algorithm, which for given disjoint sets $X, Z \subset N$, indicates those nodes $y \in N$ which are not c-separated by Z from X [27]. However, in general, we do not know what is the complexity of a testing algorithm based on the separation criterion in comparison with an algorithm based on the moralization criterion. Our naive estimate is that the separation criterion can appear to be more suitable in the case where one has to read a *dependency statement* from a CG (since one needs to find just one active route) while in the case where one has to confirm an independency statement represented in a CG the moralization criterion should be more suitable.

In [10] the concept of d-separation was generalized to the concept of D-separation to handle more effectively the case when deterministic (i.e. functionally dependent) variables in a DAG are specified. Perhaps c-separation could be also generalized in this direction, too.

Our strong completeness result justifies completely the use of CGs in probabilistic reasoning. However, as mentioned in the remark concluding Subsection 7.1, the question of the existence of a probability distribution given a prescribed CG-model and domains for variables remains open. We share the reviewer's opinion that such a (perfectly Markovian) distribution exists in the case of non-trivial (i.e. two-element) prescribed domains. We hope that the method used by Meek [21] in the case of DAGs can be extended to the case of CGs. In fact, Meek showed that under a suitable parametrization the class of non-perfectly Markovian distributions with prescribed domains becomes a lower-dimensional variate in the class of all Markovian distributions with prescribed domains. Thus, the main problem seems to be to find an analogous parametrization suitable for every CG (perhaps the parametrization method used in Section 5, for special simple CGs only, can be generalized for this purpose).

Another possible direction of future research activity is to extend the completeness result to more general graphs like, for example, Koster's *reciprocal* *graphs* [14]. We conjecture that our completeness result can be proved in the framework of general hybrid graphs provided that the corresponding concept (like Markovian distribution etc.) can be extended to that case. Our conjecture is inspired by Spirtes [24] who extended the completeness result for DAGs [11] to the case of general directed graphs without problems.

A further open question is whether an analogous completeness result can be proved also in framework of other (non-probabilistic) calculi for dealing with uncertainty in artificial intelligence [25].

From the point of view of the calculus of relational databases and for the class of all probability distributions (i.e. not only strictly positive ones), the semi-graphoid properties are of more interest than the graphoid properties. Namely, if a functional dependency among the variables occurs, then the distribution is not strictly positive and the intersection property cannot be applied. Therefore, it would be interesting to find (for every CG) an alternative list of independency statements such that the semigraphoid closure of the list is the corresponding CG-model and that the list has minimal cardinality within the class of such semigraphoid-generating lists. Of course, as mentioned in the remark concluding Section 3 the number of items of such a list can to exceed the number of nodes in the CG in general.

In [4] criteria for reading conditional *dependence* statements from a graph which is a minimal I-map [22] of some graphoid were developed (both in the case of UGs and DAGs). Is it possible to develop a similar criterion for CGs which preferably generalizes these two graphical criteria?

In [22] UG-models were characterized as independency models satisfying certain properties (axioms), analogous to the graphoid properties. Similar properties for DAG-models were formulated there too, although they do not characterize DAG-models completely. It would be interesting to find such an axiomatic characterization of CG-models.

Acknowledgments. We thank the reviewers for their helpful comments, Dan Geiger for his constructive discussion, Nic Wilson for help with terminology, Václav Kelar for preparation of special TEX fonts, Thomas Richardson for correction of (not only) grammatical lapses, and František Matúš for advice leading to the example from Figure 5. Our thanks are also devoted to the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, UK. The revised version of this paper was prepared while the first author was a visitor there.

REFERENCES

[1] Andersson, S. A., Madigan, D. and Perlman, M. D. An alternative Markov property for chain graphs. *Uncertainty in Artificial Intelligence 12* (E. Horvitz, F. Jensen eds.) 40–48. Morgan Kaufmann 1996.

[2] Andersson, S. A., Madigan, D. and Perlman, M. D. On the Markov equivalence of chain graphs, undirected graphs and acyclic digraphs. *Scand. J. Statist.* **24** (1997) 81–102.

[3] Andersson, S. A., Madigan, D. and Perlman, M. D. A characterization of Markov equivalence classes for acyclic digraphs. *Ann. Statist.* **25** (1997) 505–541.

[4] BOUCKAERT, R. R. Bayesian networks: from construction to inference. Ph.D. dissertation, University of Utrecht 1995.

[5] BOUCKAERT, R. R. AND STUDENÝ, M. Chain graphs: semantics and expressiveness. *Symbolic and Quantitative Approaches to Reasoning and Uncertainty* (Ch. Froidevaux, J. Kohlas eds.) 67–76. Lecture Notes in Artificial Intelligence 946, Springer-Verlag 1995.

[6] BUNTINE, W. L. Chain graphs for learning. *Uncertainty in Artificial Intelligence 11* (P. Besnard, S. Hanks eds.) 46–54. Morgan Kaufmann 1995.

[7] COX, D. R. AND WERMUTH, N. *Multivariate Dependencies: Models Analyses and Interpetation.* Chapman & Hall, London 1996.

[8] DAWID, A. P. Conditional independence in statistical theory. *J. Roy. Statist. Soc. ser. B* **41** (1979) 1–31.

[9] FRYDENBERG, M. The chain graph Markov property. *Scand. J. Statist.* **17** (1990) 333–353.

[10] GEIGER, D., VERMA, T. S. AND PEARL, J. Identifying independence in Bayesian networks. *Networks* **20** (1990) 507–534.

[11] GEIGER, D. AND PEARL, J. On the logic of causal models. *Uncertainty in Artificial Intelligence 4* (R. D. Shachter, T. S. Lewitt, L. N. Kanal, J. F. Lemmer eds.) 3–14. North-Holland 1990.

[12] GEIGER, D. AND PEARL, J. Logical and algorithmic properties of conditional independence and graphical models. *Ann. Statist.* **21** (1993) 2001–2021.

[13] KIIVERI, H., SPEED, T. P. AND CARLIN, J. B. Recursive causal models. *J. Aust. Math. Soc. ser. A* **36** (1984) 30–52.

[14] KOSTER, J. T. A. Markov properties of non-recursive causal graphs. *Ann. Statist.* **24** (1996) 2148–2177.

[15] LAURITZEN, S. L. AND WERMUTH, N. Mixed interaction models. Res. rep. R-84-8, Inst. Elec. Sys., University of Aalborg 1984 (the research report was later modified and became a basis of the journal paper [16]).

[16] LAURITZEN, S. L. AND WERMUTH, N. Graphical models for associations between variables, some of which are qualitative and some quantitative. *Ann. Statist.* **17** (1989) 31–57.

[17] LAURITZEN, S. L. Mixed graphical association models. *Scand. J. Statist.* **16** (1989) 273–306.

[18] LAURITZEN, S. L., DAWID, A. P., LARSEN, B. N. AND LEIMER, H.-G. Independence properties of directed Markov fields. *Networks* **20** (1990) 491–505.

[19] LAURITZEN, S. L. *Graphical Models.* Clanderon Press, Oxford 1996.

[20] MEEK, C. Causal inference and causal explanation with background knowlewdge. *Uncertainty in Artificial Intelligence 11* (P. Besnard, S. Hanks eds.) 403–410. Morgan Kaufmann 1995.

[21] MEEK, C. Strong completeness and faithfulness in Bayesian networks. *Uncertainty in Artificial Intelligence 11* (P. Besnard, S. Hanks eds.) 411–418. Morgan Kaufmann 1995.

[22] PEARL, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann 1988.

[23] RICHARDSON, T. S. Chain graphs and symmetric associations. *Learning in Graphical Models* (M. I. Jordan ed.) 231–260. Kluwer, Dordrecht 1998.

[24] SPIRTES, P. Directed cyclic graphical representations of feedback models. *Uncertainty in Artificial Intelligence 12* (E. Horvitz, F. Jensen eds.) 454–461. Morgan Kaufmann 1996.

[25] STUDENÝ, M. Formal properties of conditional independence in different calculi of AI. *Symbolic and Quantitative Approaches to Reasoning and Uncetainty* (M. Clarke, R. Kruse, S. Moral eds.) 341–348. Lecture Notes in Computer Science 747, Springer-Verlag 1993.

[26] STUDENÝ, M. A recovery algorithm for chain graphs. *Int. J. Approx. Reasoning* **17** (1997) 265–293.

[27] Studený, M. Bayesian networks from the point of view of chain graphs. To appear in *Uncertainty in Artificial Intelligence 14* (G. Gooper, S. Moral eds.). Morgan Kaufmann 1998.

[28] Verma, T. S and Pearl, J. Causal netwoks: semantics and expressiveness. *Uncertainty in Artificial Intelligence 4* (R. D. Shachter, T. S. Lewitt, L. N. Kanal, J. F. Lemmer eds.) 69–76. North-Holland 1990.

[29] Volf, M. and Studený, M. A graphical characterization of the largest chain graphs. In preparation.

[30] Wermuth, N. and Lauritzen, S. L. On substantive research hypotheses, conditional independence graphs and graphical chain models. *J. Roy. Statist. Soc. ser. B* **52** (1990) 21–50.

[31] Whittaker, J. *Graphical Models in Aplied Multivariate Statistics.* John Wiley 1990.

Inst. of Inform. Theory & Autom.
Academy of Sciences of Czech Rep.
Prague, Czech Republic, CZ18208
studeny@utia.cas.cz

Dept. of Computer Science
Utrecht University
Utrecht, The Netherlands, 3508 TB
rrb@xtra.co.nz

Lab. of Intel. Systems
University of Economics
Prague, Czech Republic, CZ14800.