Qualitative Chain Graphs and their Use in Medicine

Martijn Lappenschaar, Arjen Hommersom, and Peter J.F. Lucas
Institute for Computing and Information Sciences,
Radboud University Nijmegen, The Netherlands
Email: {mlappens, arjenh, peterl}@cs.ru.nl

Abstract

For modelling diseases in medicine, chain graphs are more attractive than directed graphs, i.e., Bayesian networks, as they support representing interactions between diseases that have no natural direction. In particular, representation by chain graphs is preferred over Bayesian networks as they have the ability to capture equilibrium models. Using qualitative abstractions of probabilistic interactions is also of interest in this context, as these would allow focusing on patterns in the interactions rather than looking at the numerical detail, which for medical purposes is of paramount importance. So far, qualitative abstractions of probabilistic interactions have been developed only for Bayesian networks in the form of the framework of qualitative probabilistic networks. In this paper, qualitative abstractions are developed for chain graphs with the practical purpose of using these as constraints on the hyperspace of probability distributions. The usefulness of this approach is explored for disease modelling.

1 Introduction

In many application fields, probabilistic graphical models are seen as convenient and intuitive formalisms to capture the probabilistic independence information of a domain. Popular graphical models include undirected graphs (UGs), also called Markov networks, and acyclic directed graphs (ADGs), also called Bayesian networks (Pearl, 1988). However, for both undirected and directed graphs one meets undesirable limitations when representing independence information for an actual problem, such as from medicine. Hybrid graphs, such as chain graphs (Lauritzen and Wermuth, 1989), that contain both directed and undirected arcs offer an elegant generalisation of both Markov networks and Bayesian networks. Chain graphs have been shown to model equilibrium systems (Lauritzen and Richardson, 2002), which occur in many areas including biology, physics, chemistry, and economics. For example, human physiology contains many regulatory mechanism that ensure homeostasis, i.e., a state of equilibrium. In particular, when modelling such processes, even when ignoring the dimension of time, Bayesian networks are not entirely suitable and more expressive models are required.

However, Bayesian networks have the advantage that both structure and parameters can be assessed from either expert knowledge, data, or both, which renders Bayesian networks whitebox rather than blackbox models. Qualitative abstractions of Bayesian networks, qualitative probabilistic networks (QPNs), provide a useful method for exploiting qualitative constraints in assessing probabilistic information. For the more expressive chain graphs, it is much more difficult to exploit human knowledge in assessing their parameters, and, as a consequence, these models are at the moment less whitebox than Bayesian networks. The aim of the research described in this paper is to come up with ways to make chain graphs more suitable as whitebox models in particular by the use of qualitative probabilistic abstractions.

While it is well known that QPN theory has its limitation when it comes to qualitative reasoning, the main reason why QPN theory is not used in actual systems, QPNs may be quite use-

ful when looked at as offering constraints when estimating a probability distribution. This, for example, allows deriving distributions over arbitrary marginals, i.e., second-order distributions (Druzdzel and van der Gaag, 1995). If an exact probability is not required, then such distributions provide insight into the domain and could, e.g., be used to make decisions.

In this paper, we will first argue why chain graphs provide a good starting point for modelling in medicine. To support a qualitative modelling approach, we will give a formal extension of QPNs based on chain graphs. We show its usefulness by semi-qualitative reasoning in a medical example, although it can be applied to any field involving a model that is represented as a chain graph.

2 Motivation from the medical field

As stated in the introduction, many physiological processes within the human body can be seen as causal feedback systems, in which some kind of equilibrium setpoint is maintained. In non-healthy people the equilibrium setpoint typically differs from the healthy people, but therapeutic interventions can reset the equilibrium setpoint to a state that is closer to the healthy people. The disturbance of the equilibrium of one physiological process, might also alter the equilibrium setpoints of other regulation systems, which might in turn induce new pathophysiology and decrease the patient's prognosis even further. In the interest of the physician, it is important to know the qualitative dynamics of such interactions, i.e., is it more likely that a therapy for a specific disease give rise to symptoms of another (patho)physiological process. Some of these aspects are illustrated in the following example, which will be used as a running example throughout the paper.

Example 1. Figure 1 shows an abstraction of the interaction between two diseases, i.e., diabetes mellitus and lipid disorder, along with its typical blood measurements, a risk factor, i.e., obesity, and a possible therapy for diabetes. It is assumed that there is feedback between the pathophysiology of both diseases, which is almost always in some kind of equilibrium. The association between these pathophysiologies can be measured by the fact that the symptoms are associated, i.e., elevated glucose levels are associated with elevated cholesterol levels.

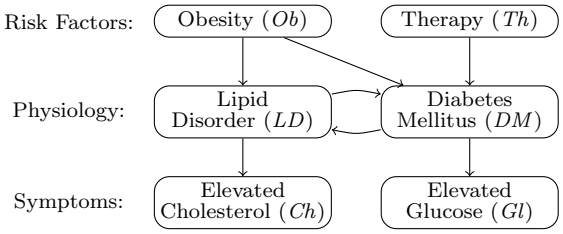

Figure 1: Schematic representation of an interaction between diabetes mellitus and a lipid disorder, showing that between the diseases feedback exists within their pathophysiology.

3 Preliminaries

3.1 Chain Graphs

A chain graph (CG) is a probabilistic graphical model that consists of labelled vertices, representing random variables, connected by directed and undirected edges. The definitions here are in accordance with existing literature on chain graphs (Studený and Bouckaert, 1998).

Let $G = (V, E)$ be a *hybrid graph*, where V denotes the set of *vertices* and E the set of *edges*, where an edge is either an *arc* (directed edge), or a *line* (undirected edge). Let indexed letters, e.g., V_1 and V_2, indicate vertices of a chain graph. We denote an arc connecting two vertices by '$\rightarrow$' and a line by '$-$'. Consider two vertices V_1 and V_2. If $V_1 \rightarrow V_2$ then V_1 is a *parent* of V_2. If $V_1 - V_2$ then V_1 is a *neighbour* of V_2. The set of parents and neighbours of a vertex V_i are denoted by $\mathrm{pa}(V_i)$ and $\mathrm{ne}(V_i)$, respectively. The set $\mathrm{pa}(V_i) \cup \mathrm{ne}(V_i)$ is the *boundary* of V_i, denoted by $\mathrm{bd}(V_i)$. We will denote $\mathrm{cl}(V_i)$ as the *closure* of V_i defined by $\mathrm{bd}(V_i) \cup \{V_i\}$.

A *path* of length n in a hybrid graph G is a sequence of distinct vertices $V_1, \ldots, V_{n+1}$, such that either $V_i - V_{i+1} \in E$, $V_i \rightarrow V_{i+1} \in E$, or $V_i \leftarrow V_{i+1} \in E$. A *directed path* is a path which includes at least one arc, and where all arcs have the same direction. A *chain graph* is a hybrid graph with the restriction that no

directed cycles exist. A *descending path* is a path, where there are no $V_i \leftarrow V_{i+1} \in E$. A vertex V_i is an ancestor of V_j if there exists an descending path from V_i to V_j. The set $\text{an}(V_i)$ denotes the set of *ancestors* of vertices from V_i.

If there is a line between every pair of vertices in a set of vertices, then this set is named *complete*. A *clique* is a maximally complete subset. Removing all the arcs from the graph leaves us with vertices connected by lines, called *chain components*; the set of all chain components is denoted here by $\mathcal{C}$. The *family* of a vertex V_i, denoted by $\text{fa}(V_i)$, is the set $C \cup \text{pa}(C)$ where $C \in \mathcal{C}$ and $V_i \in C$.

Associated to a chain graph $G = (V, E)$ is a joint probability distribution over the set of vertices $P(V)$ that is faithful to the chain graph G, i.e., it includes all the independencies implied by the graph. In this paper, we assume $P(V)$ to be a strictly positive discrete factorisable distribution, which are almost always faithful (Peña, 2009), defined by an *outer factorisation*:

$$P(V) = \prod_{C \in \mathcal{C}} P(C \mid \text{pa}(C)) \tag{1}$$

where each $P(C \mid \text{pa}(C))$ is defined by a clique-wise factorisation:

$$P(C \mid \text{pa}(C)) = Z^{-1}(\text{pa}(C)) \prod_{M \in M_C} \varphi_M(M) \tag{2}$$

given that M_C are the complete subsets in the closure graph of C, i.e., the subgraph $G_{C \cup \text{pa}(C)}$ where each arc is replaced by a line and each distinct vertex of $\text{pa}(C)$ is also connected by a line. The functions φ are non-negative real functions, called *potentials*; they generalise joint probability distributions in the sense that they do not need to be normalised. The constant $Z(\text{pa}(C)) = \sum_C \prod_{M \in M_C} \varphi_M(M)$ normalises the product to a probability distribution.

Undirected edges in chain graphs can be interpreted as an equilibrium (steady-state) in a feedback model (Lauritzen and Richardson, 2002). For example, consider again the graph in Figure 1, where there is a feedback relationship between lipid disorder and diabetes mellitus. In practice, this feedback system is in a steady state, although the setpoint of the feedback system may be changed, for example the amount of insulin resistance. Therefore, only the relationships between variables within a steady-state are relevant, rather than the underlying dynamic process that leads to the equilibrium. Moreover, the underlying dynamics is very difficult to measure *in vivo*, hence, the parameters of such models are difficult to elicit. Therefore, we argue that chain graph offer an attractive abstraction of the underlying dynamic mechanism for feedback systems in disease models. The corresponding chain graph and factorisation of Figure 1 in its steady state is shown in Figure 2.

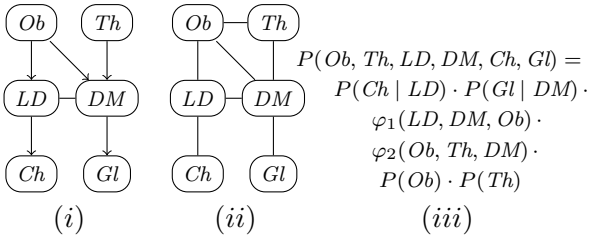

Figure 2: Chain graph representation (i), closure graph of chain components (ii), and factorisation (iii) of the example in Figure 1.

3.2 QPNs

Qualitative probabilistic networks (QPNs) were introduced by Wellman (1990), as a qualitative abstraction of Bayesian networks. Conditional probability distributions are replaced by qualitative knowledge in the form of signs, which describes the relationships among variables by the concepts of probabilistic influences and synergies. Here we briefly recall the theory in accordance with the definitions of Renooij (2001).

For clarity of exposition, we will assume that each node $A \in V$ in the discrete chain graph model is a binary variable, which can take the values a ($A = \text{true}$) and $\overline{a}$ ($A = \text{false}$). Further, for notational convenience, we will sometimes write the singleton set $\{A\}$ as A, and, if $X, Y \subseteq V$, then we will write XY instead of $X \cup Y$. Finally, we denote $X - Y$ for $X \setminus Y$. For example $X - AB$ is an abbreviation of $X \setminus \{A, B\}$.

Then, a *qualitative influence* expresses how the value of one variable influences the probabil-

ity of observing values of another variable. Let Z denotes the set of variables $\text{pa}(B) - A$. We say that A has a *positive qualitative influence on* B, if $P(b \mid a, z) - P(b \mid \overline{a}, z) \geq 0$, regardless of the configuration z, with a strict inequality for at least one configuration z.

An *additive synergy* expresses how the interaction between two variables influences the probability of observing the values of a third variable. Now, let Z denotes the set consisting the variables $\text{pa}(B) - A_1A_2$. We say there is a *positive additive synergy of* A_1 *and* A_2 *on* B, if $P(b \mid a_1, a_2, z) + P(b \mid \overline{a}_1, \overline{a}_2, z) - P(b \mid \overline{a}_1, a_2, z) - P(b \mid a_1, \overline{a}_2, z) \geq 0$, regardless of the configuration z, with a strict inequality for at least one configuration z.

A *product synergy* expresses how upon observation of a common child of two vertices, observing the value of one parent vertex influences the probability of observing a value of the other parent. We say there is a *positive product synergy* of A_1 and A_2 with regard to the value b on variable B, if $P\big(b \mid a_1, a_2, z\big) \cdot P\big(b \mid \overline{a}_1, \overline{a}_2, z\big) - P\big(b \mid \overline{a}_1, a_2, z\big) \cdot P\big(b \mid a_1, \overline{a}_2, z\big) \geq 0$, regardless of the configuration z, with a strict inequality for at least one configuration z.

Negative and *zero* influences and synergies are defined analogously, by replacing $\geq$ with $\leq$ and $=$ respectively. If none of these cases hold, we say that the influence or synergy is *ambiguous*.

4 Qualitative Chain Graphs

In this section, we will analyse influences and synergies in the context of chain graph models.

4.1 Influences in chain graphs

The properties of signs in qualitative probabilistic networks rely on the fact that signs hold in any context, i.e., intuitively, a variable A positively influences another variable B if in any possible context the probability of B is higher for a compared to $\overline{a}$. While such a context is relatively clear in case of directed arcs, it is more subtle for probabilistic chain graphs, in which influences can also exist through lines. To obtain a proper definition in such a network, we will define influences in terms of *interventions* (Lauritzen and Richardson, 2002) on particular variables in the chain graph.

Definition 1. The influence of A on B in a context $c \in V - AB$, where A and B are two vertices, is the probability $P(b \mid\mid a, c) - P(b \mid\mid \overline{a}, c)$ where $P(B \mid\mid A, C)$ is the probability of B after an intervention on A and C.

We say that A has a positive influence on B if the influence of A on B is ≥ 0 in any context. Negative, zero and ambiguous influences are defined similarly. Similarly to QPNs, these influences coincide with a difference in conditional probabilities if we assume that the chain graph can be given a causal interpretation and the chain components model equilibria, which will be assumed in the remainder of this paper.

Lemma 1. *If a chain graph $G = (V, E)$ is generated by a causal feedback model where lines represent equilibria (Lauritzen and Richardson, 2002), and $P(B \mid\mid V - B)$ denotes the probability distribution of B after an intervention on all other variables, then:*

$$P(B \mid\mid V - B) = P(B \mid \text{fa}(B) - B)$$

Proof. This is a direct corollary of Equation (18) in Lauritzen and Richardson (2002). □

Given the fact that a node is independent of its non-descendants given its boundary, i.e., for models that factorise as given in Section 3.1, a local Markov property holds, which yields the following lemma.

Lemma 2. *Given a chain graph $G = (V, E)$ and vertices $A, B \in V$ such that $A \in \text{bd}(B)$, then $P(B \mid \text{an}(B) - B) = P(B \mid \text{fa}(B) - B) = P(B \mid \text{bd}(B))$.*

Proof. Follows from the local Markov property of chain graphs (Frydenberg, 1990): $V_i \perp\!\!\!\perp \text{an}(V_i) - \text{cl}(V_i) \mid \text{bd}(V_i)$. □

Then, using Lemmas 1 and 2, we obtain an expression of influences in chain graphs in terms of conditional probabilities.

Proposition 1. *Given two nodes A and B and a context c, then the influence of A on B in context c equals:*

$$P(b \mid a, z) - P(b \mid \overline{a}, z)$$

where $c = z \cup x$, $Z = \mathrm{bd}(B) - A$, *and* $X = V - ZAB$.

Note that it follows that the influence of node A on node B is a zero influence if $A \notin \mathrm{bd}(B)$ and $A \in \mathrm{an}(B)$. Also note that the qualitative influences generalise the QPN definitions, since $\mathrm{bd}(B) = \mathrm{pa}(B)$ for any $B \in V$ if every chain components consist of a single vertex.

4.2 Symmetry of influences

In QPNs, the qualitative signs are symmetric, i.e., if there is some influence from a node A to a node B, then there is an influence from B to A with the same sign. Therefore, only a single sign is needed for every arc in a QPN. In the remainder, we will prove that this symmetry is preserved for qualitative chain graphs, i.e., also for neighbouring nodes the signs are symmetric. First we prove a lemma that rephrases qualitative influences in terms of relationships between potential functions. For notational convenience, we do not write universally quantified variables in potentials, e.g., $\varphi_M(a)$ is shorthand for $\varphi_M(M - A, a)$, and we will write $\varphi_M(a, X)$ for $\varphi_M(X)$ if $A \notin M$ for any $X \subseteq V$. Further, we will focus in the next lemma and theorem on positive influences, however, the same reasoning holds for negative and zero influences.

Lemma 3. *Given a chain graph G containing vertices A and B, with $A \in \mathrm{bd}(B)$ and B an element of a component C, it holds that:*

$$P(b \mid a, \mathrm{fa}(B) - AB) \geq P(b \mid \overline{a}, \mathrm{fa}(B) - AB)$$

if and only if

$$\prod_{M \in M_{AB}} \varphi_M(a, b)\varphi_M(\overline{a}, \overline{b}) \geq \prod_{M \in M_{AB}} \varphi_M(a, \overline{b})\varphi_M(\overline{a}, b)$$

where $M_{AB} = \{M \in M_C \mid \{A, B\} \subseteq M\}$.

Proof. By basic probability theory, we have:

$$P(B \mid A, \mathrm{fa}(B) - AB) = \frac{P(C)}{P(C - B)} = \frac{P(C \mid \mathrm{pa}(C))P(\mathrm{pa}(C))}{\sum_B P(C \mid \mathrm{pa}(C))P(\mathrm{pa}(C))}$$

Using Equation (2), that factorises conditional probabilities of a component into potentials, the left-hand side therefore equals to:

$$\frac{Z^{-1}(\mathrm{pa}(C))\Big(\prod_{M_C} \varphi_M(a, b)\Big)P(\mathrm{pa}(C))}{\sum_B Z^{-1}(\mathrm{pa}(C))\Big(\prod_{M_C} \varphi_M(a, B)\Big)P(\mathrm{pa}(C))} \geq \frac{Z^{-1}(\mathrm{pa}(C))\prod_{M_C}\Big(\varphi_M(\overline{a}, b)\Big)P(\mathrm{pa}(C))}{\sum_B Z^{-1}(\mathrm{pa}(C))\Big(\prod_{M_C} \varphi_M(\overline{a}, B)\Big)P(\mathrm{pa}(C))}$$

Given that $B \in C$, we have $B \notin \mathrm{pa}(C)$, so the term $Z^{-1}(\mathrm{pa}(C))P(\mathrm{pa}(C))$ only depends on A. By replacing this term by $f(A)$ and multiplying each side by the denominators, we obtain:

$$\prod_{M_C} \varphi_M(a, b)f(a) \sum_B \prod_{M_C} \varphi_M(\overline{a}, B)f(\overline{a}) \geq \prod_{M_C} \varphi_M(\overline{a}, b)f(\overline{a}) \sum_B \prod_{M_C} \varphi_M(a, B)f(a)$$

Writing out the possible values for the summation over B, i.e., b and $\overline{b}$, we obtain:

$$\prod_{M_C} \varphi_M(a, b)\varphi_M(\overline{a}, b)f(a)f(\overline{a}) + \prod_{M_C} \varphi_M(a, b)\varphi_M(\overline{a}, \overline{b})f(a)f(\overline{a}) \geq \prod_{M_C} \varphi_M(\overline{a}, b)\varphi_M(a, b)f(a)f(\overline{a}) + \prod_{M_C} \varphi_M(\overline{a}, b)\varphi_M(a, \overline{b})f(a)f(\overline{a})$$

Removing the factors that are the same on both sides of the equation we get:

$$\prod_{M_C} \varphi_M(a, b)\varphi_M(\overline{a}, \overline{b}) \geq \prod_{M_C} \varphi_M(a, \overline{b})\varphi_M(\overline{a}, b)$$

For all potentials not depending on both A and B, e.g., $\varphi_M(A, B) = \varphi_M(B)$, its corresponding factors are also the same on both sides of the equation, leaving us with:

$$\prod_{M \in M_{AB}} \varphi_M(a, b)\varphi_M(\overline{a}, \overline{b}) \geq \prod_{M \in M_{AB}} \varphi_M(a, \overline{b})\varphi_M(\overline{a}, b)$$

□

	ld, dm	$\overline{ld}, dm$	$ld, \overline{dm}$	$\overline{ld}, \overline{dm}$
ob	16	4	2	4
$\overline{ob}$	2	2	1	5

Table 1: Potentials over Ob, LD, and DM.

Example 2. Continuing Example 1, to evaluate a (say, positive) influence of Ob on LD only involves φ_1 (cf. Figure 1c). Therefore, a positive influence of Ob on LD is equivalent to:

$$\varphi_1(ob, ld, DM)\varphi_1(\overline{ob}, \overline{ld}, DM) \geq \varphi_1(\overline{ob}, ld, DM)\varphi_1(ob, \overline{ld}, DM)$$

for all values of DM. Consider, as an example, the potential defined in Table 1. It holds for $dm \rightarrow 16 \geq 4$, and for $\overline{dm} \rightarrow 20 \geq 8$, implying a positive influence of Ob on LD. Indeed, by computing the individual probabilities using Table 1, we obtain:

$$P(ld \mid ob, dm) = 8/9 \geq P(ld \mid \overline{ob}, dm) = 2/3$$
$$P(ld \mid ob, \overline{dm}) = 2/3 \geq P(ld \mid \overline{ob}, \overline{dm}) = 4/9$$

As a result of this lemma, determining the nature of a qualitative influence between two vertices implies that one only has to consider those potentials for cliques containing the two variables that describe the influence. Hence, we have the following result that we were aiming for, proving the symmetry between qualitative influences between arbitrary edges.

Theorem 1. *It holds that qualitative signs of chain graphs are symmetric, i.e., suppose* $(A, B) \in E$*, then* $P(b \mid a, X) - P(b \mid \overline{a}, X) \geq 0$ *if and only if* $P(a \mid b, Y) - P(a \mid \overline{b}, Y) \geq 0$*, where* $X = \mathrm{bd}(B) - A$ *and* $Y = \mathrm{bd}(A) - B$*.*

Proof. ($\Rightarrow$) Assume $P(b \mid a, X) - P(b \mid \overline{a}, X) \geq 0$. Since $Y \subseteq \mathrm{an}(B) - \mathrm{cl}(B)$, by the local Markov property, we have $B \perp\!\!\!\perp Y \mid AX$, so it follows that $P(b \mid a, X, Y) - P(b \mid \overline{a}, X, Y) \geq 0$. By Proposition 1 and Lemma 3, it follows that

$$\prod_{M \in M_{AB}} \varphi_M(a, b)\varphi_M(\overline{a}, \overline{b}) \geq \prod_{M \in M_{AB}} \varphi_M(a, \overline{b})\varphi_M(\overline{a}, b).$$

By the same reasoning, we get $P(a \mid b, x, y) - P(a \mid \overline{b}, x, y) \geq 0$ and given that $A \perp\!\!\!\perp X \mid BY$, we conclude $P(a \mid b, y) - P(a \mid \overline{b}, y) \geq 0$. The converse holds by the same steps. □

4.3 Qualitative chain graphs defined

Given the properties of influences in chain graphs, we are now in the position to define the usual notions of qualitative probabilistic networks for chain graphs. We will focus on the positive influences and synergies; the negative, zero and ambiguous influences and synergies are defined similarly.

Definition 2. We say that a vertex A positively influences a vertex B, written as $S^+(A, B)$, iff $A \in \mathrm{bd}(B)$ and

$$P(b \mid a, \mathrm{bd}(B) - A) \geq P(b \mid \overline{a}, \mathrm{bd}(B) - A)$$

A positive *additive synergy* expresses that the joint influence of A_1 and A_2 is greater (or less in case of a negative synergy) than their separate influence on a child B. More formally, we say that influences of A_1 on B are higher in contexts c where $a_2 \in c$ compared to contexts c' where $\overline{a}_2 \in c$. This can be rephrased in terms of conditional probabilities following Proposition 1.

Definition 3. We say that vertices A_1 and A_2 express a positive additive synergy on a vertex B, written as $Y^+(\{A_1, A_2\}, B)$, iff $A_1, A_2 \in \mathrm{bd}(B)$, $Z = \mathrm{bd}(B) - A_1A_2$, and

$$P(b \mid a_1, a_2, Z) - P(b \mid \overline{a}_1, a_2, Z) \geq P(b \mid a_1, \overline{a}_2, Z) - P(b \mid \overline{a}_1, \overline{a}_2, Z)$$

A *product synergy* expresses how the value of one cause influences the probability of the value of another cause when observing the common child the child. By similar reasoning for influences and additive synergies, we define product synergies as follows.

Definition 4. We say that vertices A_1 and A_2 express a negative product synergy with regard to the value b on the vertex B, written as $X^+(\{A_1, A_2\}, b)$, iff $A_1, A_2 \in \mathrm{bd}(B)$, $Z = \mathrm{bd}(B) - A_1A_2$, and

$$P(b \mid a_1, a_2, Z) \cdot P(b \mid \overline{a}_1, \overline{a}_2, Z) \geq P(b \mid a_1, \overline{a}_2, Z) \cdot P(b \mid \overline{a}_1, a_2, Z)$$

5 Experimental Results

Probabilistic inference with a QPN can be done using sign-propagation, based on message-passing between neighbouring nodes (Druzdzel and Henrion, 1993), which has its limitations in case of trade-offs. An alternative approach is to look upon the qualitative signs as constraints on the joint probability distribution, as proposed in Druzdzel and van der Gaag (1995), where a canonical representation consisting of (in)equalities expressing constraints on the hyperspace of possible joint probability distributions is used. In this approach, some of the conditional probabilities or cliques may be elicited from experts or learned from data, where for others, only qualitative information is available.

In this paper, we take a similar approach, where we sample the unknown potentials from the factorisation of a given chain graph (cf. Equations (1) and (2)). Instead of sampling the full joint probability distributions and then establishing if the distribution is consistent with qualitative influences, the potentials can be sampled more efficiently by Lemma 3, as this shows that influences impose *local* constraints on the potentials. Likewise, synergies can be stated in terms of constraints on the local potentials using the following proposition.

Proposition 2. *Given a chain graph G containing vertices A_1, A_2, and B, with $A_1, A_2 \in \mathrm{bd}(B)$ and B an element of a component C, it holds that a positive additive synergy $Y^+(\{A_1, A_2\}, B)$ exists if and only if*

$$\phi_C(a_1, a_2, b) \;+\; \phi_C(\overline{a}_1, \overline{a}_2, b) \geq \phi_C(a_1, \overline{a}_2, b) \;+\; \phi_C(\overline{a}_1, a_2, b)$$

and, likewise, a positive product synergy $X^+(\{A_1, A_2\}, b)$ exists if and only if

$$\phi_C(a_1, a_2, b) \;\cdot\; \phi_C(\overline{a}_1, \overline{a}_2, b) \geq \phi_C(a_1, \overline{a}_2, b) \;\cdot\; \phi_C(\overline{a}_1, a_2, b)$$

$$\text{with } \phi_C(a_1, a_2, b) = \frac{\prod_{M \in M_{A_1 A_2 B}} \varphi(a_1, a_2, b)}{\sum_B \prod_{M \in M_{A_1 A_2 B}} \varphi(a_1, a_2, B)}.$$

Procedure 1 *sample-distribution*(potentials $\phi_{known}, \phi_{unknown}$, qualitative constraints C)

```
for φ_M ∈ φ_unknown do
    φ_M ← sample a potential for variables M
    while !satisfies[1](φ_M, C) do
        φ_M ← resample potential for M
    end while
    φ_known ← φ_known ∪ {φ_M}
end for
return distribution using Equations 1 and 2
```

The proof of Proposition 2 follows the same line as the proof of Lemma 3, but is omitted here due to lack of space.

Given these properties, distributions can be sampled that satisfy the qualitative constraints (cf. Procedure 1). Then, using these samples, second-order distributions of arbitrary marginal distributions can be derived in a straightforward manner. While typically the marginals range over the whole $[0, 1]$ interval, the qualitative constraints alter the shape (e.g., the mean and variance) of the distribution, which can then be used to draw conclusions from the model.

Example 3. Continuing from Example 1, consider the quantitative and qualitative information available in Figure 3a. The 2nd order distribution of *Ch*, i.e., high cholesterol, is then shown in Figure 3b. An intervention on *Th*, yields the 2nd order distribution in Figure 3c, showing that probability of *Ch* shifts to lower values. The probability $P(Ch \mid Th) < P(Ch) \approx 0.82$ within the generated samples (n=100.000), suggesting with high confidence that diabetic therapy is also beneficial to reduce cholesterol levels. Note that it has been derived without any quantitative information about the chain component containing *LD* and *DM*. An additional positive synergy between *Ob* and *Th* on *DM* pushes the distribution even more to lower probabilities with an intervention on *Ob*, see Figure 3d. The probability $P(Ch \mid \overline{Ob}, Th) < P(Ch) \approx 0.91$, suggesting that an additional reduction of weight in combination with diabetic therapy is even more beneficial to reduce cholesterol levels.

[1] By local constraints of Lemma 3 and Proposition 2.

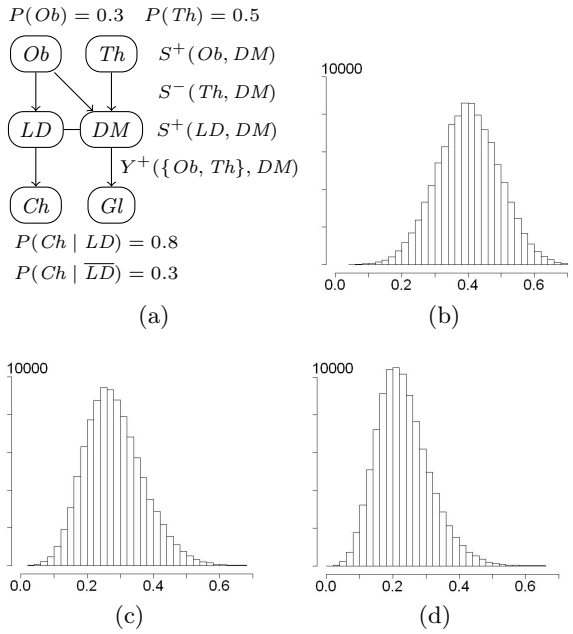

Figure 3: Qualitative and quantitative information (a) of Figure 2(i), and 2nd order probability distributions of *Ch* (b,c,d), in the presence of specific interventions (see Example 3).

6 Conclusions

In this paper we extended the QPN framework, i.e., qualitative influences and synergies, towards chain graphs. We analysed influences in chain graphs and showed some of its basis properties. This allows the modelling of qualitative representation of feedback systems, e.g., (patho)physiological processes, since chain graphs have the ability to model equilibria of such systems. We have illustrated that within the same chain graph it is feasible to combine quantitative and qualitative information.

This is of importance for medicine, i.e., without knowing the exact joint probability distribution that exists between diseases, we are still able to draw qualitative conclusions on the dynamics that exist within a disease model. Starting from quantitative information, e.g., conditional probabilities relating symptoms and diseases, adding qualitative information, can be efficiently done by putting specific constraints on the potentials of local cliques in the chain graph. Such information often comes in terms of so-called odds ratios, which easily translates to qualitative influence and synergies.

In future work we aim to apply this formalism in a study on diabetes and cardiovascular comorbidities involving multiple feedback systems. Furthermore, the sampling of potentials may be improved by exploiting Monte Carlo methods which take into account bounds on the hyperspace (e.g. based on (Smith, 1984)).

References

MJ Druzdzel and M Henrion. 1993. Efficient reasoning in qualitative probabilistic networks. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, California. AAAI Press.

MJ Druzdzel and LC van der Gaag. 1995. Elicitation of probabilities for belief networks: Combining qualitative and quantitative information. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 141–148.

M. Frydenberg. 1990. The chain graph Markov property. *Scand J Statist*, 17:333–353.

SL Lauritzen and TS Richardson. 2002. Chain graph models and their causal interpretations. *J.R. Statist. Soc. B*, 64(3):321–361.

SL Lauritzen and N Wermuth. 1989. Graphical models for associations between variables, some of which are qualitative and some quantitative. *The Annals of Statistics*, 17(1):31–57.

J Pearl. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

JM Peña. 2009. Faithfulness in chain graphs: The discrete case. *J. of Approximate Reasoning*, 50(8):1306–1313.

S Renooij. 2001. *Qualitative Approaches to Quantifying Probabilistic Network*. Ph.D. thesis, Utrecht University.

RL Smith. 1984. Efficient monte carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32(6):1296–1308.

M Studený and RR Bouckaert. 1998. On chain graph models for description of conditional independence structures. *The Annals of Statistics*, 26(4):1434–1495.

MP Wellman. 1990. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 40:257–303.