The Computational Complexity of Monotonicity in Probabilistic Networks

Johan Kwisthout*

Department of Information and Computer Sciences, University of Utrecht,
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands.
johank@cs.uu.nl

Abstract. Many computational problems related to probabilistic networks are complete for complexity classes that have few 'real world' complete problems. For example, the decision variant of the inference problem (PR) is PP-complete, the MAP-problem is NP^{PP}-complete and deciding whether a network is monotone in mode or distribution is co-NP^{PP}-complete. We take a closer look at monotonicity; more specific, the computational complexity of determining whether the values of the variables in a probabilistic network can be ordered, such that the network is monotone. We prove that this problem – which is trivially co-NP^{PP}-hard – is complete for the class co-$\text{NP}^{\text{NP}^{\text{PP}}}$ in networks which allow implicit representation.

1 Introduction

Probabilistic networks [6] (also called Bayesian or belief networks) represent a joint probability distribution on a set of statistical variables. A probabilistic network is often described by a directed acyclic graph and a set of conditional probabilities. The nodes represent the statistical variables, the arcs (or lack of them) represent (in)dependencies induced by the joint probability distribution. Probabilistic networks are often used in decision support systems such as medical diagnosis systems (see e.g. [2] or [11]). Apart from their relevance in practical situations, they are interesting from a theoretical viewpoint as well.

Many problems related to probabilistic networks happen to be complete for complexity classes that have few 'real world' complete problems. For example, the decision variant of the inference problem PR (is the probability of a specific instantiation of a variable higher than p) is PP-complete [3], where the exact inference problem is #P-complete [7]. The problem of finding the most probable explanation (MPE), i.e., the most likely instantiation to all variables, has an NP-complete decision variant [8]. On the other hand, determining whether there is an instantiation to a subset of all variables (the so-called MAP variables), such that there exists an instantiation to the other variables with probability higher

* The work of this author was partially supported by the Netherlands Organisation for Scientific Research NWO

than p (the MAP problem) is NP^{PP}-complete [5]. Determining whether a network is monotone (in mode or in distribution) is co-NP^{PP}-complete [10].

Monotonicity is often studied in the context of probabilistic classification, where a network is constructed of evidence variables (like observable symptoms and test results), non-observable intermediate variables, and one or more classification variables. Informally, the conditional probability of a variable C given evidence variables E is monotone, if higher ordered instantiations to E always lead to higher values of C (isotone) or always lead to lower values of C (antitone). The question whether these relations are monotone is particularly relevant during the construction and verification of the network. Often, domain experts will declare that certain relations ought to be monotone, and the conditional probabilities in the network should then respect these assumptions. When a violation is found, the probabilities should be reconsidered, by elicitating better estimations or using more data to learn from.

While complexity results are known for the MONOTONICITY problem when all variables have fixed orderings, no such results have been obtained yet for the related problem where no such fixed order is presumed. Nevertheless, while variables sometimes have a trivial ordering (e.g., *always* > *sometimes* > *never*), such an ordering might be arbitrary, and determining a 'good' ordering might reduce the part of the network where monotonicity is violated. This problem is interesting from a theoretical viewpoint as well. If we can determine whether adding this extra 'degree of freedom' to the MONOTONICITY problem 'lifts' the complexity of the problem into a broader class, we gain some insight in the properties and power of these types of complexity classes.

In the remainder of this paper, some preliminaries are introduced in Section 2, and various monotonicity problem variants and their computational complexity are discussed in Section 3. In Section 4, we present an (alternative) proof for a restricted version of the MONOTONICITY problem as presented in [10]. This proof technique is then used in Section 5 to show that the MONOTONICITY problem with no fixed orderings, is indeed complete for the class co-$\text{NP}^{\text{NP}^{\text{PP}}}$ if we allow a (rather broad) implicit probability representation. Finally, in Section 6 these results are discussed and the paper is concluded.

2 Preliminaries

Before formalising the problems for which we want to determine their computational complexity, we first need to introduce some definitions and notations. Let $\mathbf{B} = (G, \Gamma)$ be a probabilistic or Bayesian network where Γ, the set of conditional probability distributions, is composed of rational probabilities, and let $\Pr$ be its joint probability distribution. The conditional probability distributions in Γ can be explicit, i.e., represented with look-up tables, or implicit, i.e., represented by a polynomial time computable function. If Γ consists only of explicit distributions then $\mathbf{B}$ will be denoted as an explicit network. Let $\Omega(V)$ denote the set of values that $V \in V(G)$ can take. Vertex A is denoted as a predecessor of B if $(A, B) \in A(G)$. For a node B with predecessors $A_1, \ldots, A_n$, the *configuration template* $\mathbf{A}$ is defined as $\Omega(A_1) \times \ldots \times \Omega(A_n)$; a particular instantiation of $A_1, \ldots, A_n$ will be denoted as a *configuration* of $\mathbf{A}$.

Monotonicity can be defined as stochastical dominance (monotone in distribution) or in a modal sense (monotone in mode)[1], furthermore monotonicity can be defined on a global scale, or locally (only relations along the arcs of the network are considered). In this paper, we discuss global monotonicity in distribution only. We distinguish between weak and strong global monotonicity.

Definition 1 (global monotonicity [10]). *Let F_{Pr} be the cumulative distribution function for a node $V \in V(G)$, defined by $F_{Pr}(v) = Pr(V \leq v)$ for all $v \in \Omega(V)$. Let C be a variable of interest (e.g., the main classifier or output variable in the network), let E denote the set of observable variables, and let $\boldsymbol{E}$ be the configuration template of E. The network is strongly monotone in E, if either*

$$e \preceq e' \rightarrow F_{Pr}(c \mid e) \leq F_{Pr}(c \mid e') \text{ for all } c \in \Omega(C) \text{ and all } e, e' \in \boldsymbol{E}\text{, or}$$
$$e \preceq e' \rightarrow F_{Pr}(c \mid e) \geq F_{Pr}(c \mid e') \text{ for all } c \in \Omega(C) \text{ and all } e, e' \in \boldsymbol{E}$$

The network is weakly monotone in E, if the network is strongly monotone in $\{E_i\}$, for all variables $E_i \in E$.

Note that all networks that are strongly monotone in E are also weakly monotone, but not vice versa: whereas the strong variant assumes a partial order on all configurations of E, the weak variant allows independent isotone or antitone effects for all variables in E. Put in another way: we could make a weakly monotone network also strongly monotone by reversing the order of the values of some variables in E, such that all effects are antitone or all effects are isotone.

The above notions of monotonicity assume an implicit *ordering* on the values of the variables involved. Such an ordering is often trivial (e.g., $x > \bar{x}$ and *always* $>$ *sometimes* $>$ *never*) but sometimes it is arbitrary, like an ordering of the organs that might be affected by a disease. Nevertheless, a certain ordering is necessary to determine whether the network is monotone, or to determine which parts of the network are violating monotonicity assumptions. Thus, for nodes where no a priori ordering is given, we want to order the values of these nodes in a way that maximises the number of monotone arcs. We define the notion of an *interpretation* of V to denote a certain ordering on $\Omega(V)$, the set of values of V. Note, that the number of distinct interpretations of a node with k values equals $k!$, the number of permutations of these values.

Definition 2 (interpretation). *An interpretation of $V \in V(G)$, denoted I_V, is a total ordering on $\Omega(V)$. We will often omit the subscript if no confusion is*

[1] For variable set E, with value assignments e and e' ($e \prec e'$) and output C, the network is isotone in distribution if $\Pr(C \mid e)$ is stochastically dominant over $\Pr(C \mid e')$. The network is isotone in mode if the most probable $c \in C$ given assignment e is lower ordered than the most probable $c \in C$ given assignment e'

possible; for arbitrary interpretations we will often use σ and τ. The interpretation set $\mathbf{I_V}$ is defined as the set of all possible interpretations of V.

In the remainder, we assume that the reader is familiar with basic concepts of computational complexity theory, such as the classes P, NP and co-NP, hardness, completeness, oracles, and the polynomial hierarchy (PH). For a thorough introduction to these subjects, we refer to textbooks like [1] and [4].

In addition to these concepts, we use the *counting hierarchy* (CH) [12, 9]. The counting hierarchy closely resembles (in fact, *contains*) the polynomial hierarchy, but also involves the class PP (probabilistic polynomial time), i.e., the class that contains languages accepted by a non-deterministic Turing Machine where the *majority* of the paths accept a string if and only if it is in that language. Recall that the polynomial hierarchy can be characterised by alternating existential and universal operators applied to P, where $\exists^P P$ equals $\Sigma_1^p = \text{NP}$, $\forall^P P$ equals $\Pi_1^p = \text{co-NP}$, while $\forall^P \exists^P \forall^P \ldots P$ equals Π_k^p and $\exists^P \forall^P \exists^P \ldots P$ equals Σ_k^p, where k denotes the number of alternating quantifiers.

A convenient way to relate the counting hierarchy to the polynomial hierarchy is by introducing an additional operator $\mathcal{C}$, where $\mathcal{C}_0^p$ equals P, $\mathcal{C}_1^p$ equals PP, and in general $\mathcal{C}_{k+1}^p = \mathcal{C} \cdot \mathcal{C}_k^p = (\mathcal{C}_k^p)^{\text{PP}}$. Interesting complexity classes can be defined using these operators $\exists^P$, $\forall^P$ and $\mathcal{C}$ in various combinations. For example, $\exists^P \mathcal{C} P$ equals the class NP^{PP}, $\forall^P \mathcal{C} P$ equals co-$\text{NP}^{\text{PP}}$and $\exists^P \forall^P \mathcal{C} P$ equals $\text{NP}^{\text{NP}^{\text{PP}}}$. Default complete problems for these kind of complexity classes are defined by Wagner [12] using quantified satisfiability variants. In this paper we consider in particular the complete problems MAJSAT, E-MAJSAT, A-MAJSAT, EA-MAJSAT and AE-MAJSAT which will be used in the hardness proofs. These problems are proven complete by Wagner [12] for the classes PP, NP^{PP}, co-NP^{PP}, $\text{NP}^{\text{NP}^{\text{PP}}}$ and co-$\text{NP}^{\text{NP}^{\text{PP}}}$, respectively. In all problems, we consider a boolean formula ϕ with n variables X_i, with $1 \leq i \leq n$, and we introduce quantifiers to bound subsets of these variables.

MAJSAT
Instance: Let $\mathbf{X}$ denote the configuration template for ϕ.
Question: Does at least half of the instantiations of $\mathbf{X}$ satisfy ϕ?

E-MAJSAT
Instance: Let $1 \leq k \leq n$, let $\mathbf{X_E}$ denote the configuration template for the variables X_1 to X_k and let $\mathbf{X_M}$ denote the configuration template for X_{k+1} to X_n.
Question: Is there an instantiation to $\mathbf{X_E}$, such that at least half of the instantiations of $\mathbf{X_M}$ satisfy ϕ?

A-MAJSAT
Instance: Sets $\mathbf{X_A}$ and $\mathbf{X_M}$ as in E-MAJSAT.
Question: Does, for every possible instantiation to $\mathbf{X_A}$, at least half of

the instantiations of $\mathbf{X_M}$ satisfy ϕ?

EA-MAJSAT

Instance: Let $1 \leq k \leq l \leq n$, let $\mathbf{X_E}$, $\mathbf{X_A}$, and $\mathbf{X_M}$ denote the configuration templates for the variables X_1 to X_k, X_{k+1} to X_l, and X_{l+1} to X_n, respectively.
Question: Is there an instantiation to $\mathbf{X_E}$, such that, for every possible instantiation of $\mathbf{X_A}$, at least half of the instantiations of $\mathbf{X_M}$ satisfy ϕ?

AE-MAJSAT

Instance: Sets $\mathbf{X_A}$, $\mathbf{X_E}$, and $\mathbf{X_M}$ as in EA-MAJSAT.
Question: Is there, for all instantiations to $\mathbf{X_A}$, a possible instantiation of $\mathbf{X_E}$, such that at least half of the instantiations of $\mathbf{X_M}$ satisfy ϕ?

In the remainder, we denote the complement of a problem P as NOT-P, with 'yes' and 'no' answers reversed with respect to the original problem P. Note that, by definition, if P is in complexity class C, then NOT-P is in co-C, and, likewise, if NOT-P is in C, then P is in co-C.

3 Monotonicity variants and their complexity

In this paper, we study the computational complexity of various variants of global monotonicity. The following problems are defined on a probabilistic network $\mathbf{B} = (G, \Gamma)$, where $G = (V, A)$ is a directed acyclic graph.

1. The STRONG GLOBAL MONOTONICITY problem is the problem of testing whether $\mathbf{B}$ is strongly globally monotone, given an interpretation for V. This problem is co-NP^{PP}-complete [10] for explicit networks.
2. The WEAK GLOBAL MONOTONICITY problem is the problem of testing whether $\mathbf{B}$ is weakly globally monotone, given an interpretation for V.
3. The GLOBAL E-MONOTONICITY problem is the problem of testing whether there exists an interpretation to $\Omega(V)$, such that $\mathbf{B}$ is globally monotone.

Note that, if there exists an interpretation such that $\mathbf{B}$ is *weakly* monotone, there must also be an interpretation such that $\mathbf{B}$ is *strongly* monotone.

WEAK GLOBAL MONOTONICITY and GLOBAL E-MONOTONICITY will be discussed in Sections 4 and 5. In these sections, we use a proof technique introduced by Park and Darwiche [5] to construct a probabilistic network $\mathbf{B}_\phi$ from a given Boolean formula ϕ with n variables. For all variables $X_i (1 \leq i \leq n)$ in this formula, we create a variable X_i in G, with possible values T and F and uniform probability distribution. For each logical operator in ϕ, we create an additional variable, whose parents are the corresponding sub-formulas (or single variable in case of a negation operator) and whose conditional probability table is equal to the truth table of that operator. For example, the $\wedge$-operator would have a conditional probability of 1 if and only if both its parents have the value T, and 0 otherwise. Furthermore, we denote the top-level operator in ϕ with V_ϕ.

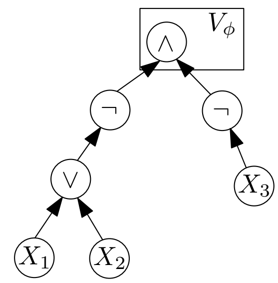

Fig. 1. The probabilistic network corresponding to $\neg(x_1 \vee x_2) \wedge \neg x_3$

In Figure 1 such a network is constructed for the formula $\neg(x_1 \vee x_2) \wedge \neg x_3$. Now, for any particular instantiation $\mathbf{x}$ of the set of all variables $\mathbf{X}$ in the formula we have that the probability of V_ϕ, given the corresponding configuration equals 1 if $\mathbf{x}$ satisfies ϕ, and 0 if $\mathbf{x}$ does not satisfy ϕ. Without any instantiation, the probability of V_ϕ is $\frac{\#_q}{2^n}$, where $\#_q$ is the number of satisfying instantiations of $\mathbf{X}$. Using such constructs, Park and Darwiche proved that the decision variant of the MAP problem is NP^{PP}-complete; we will use this construct as a starting point to prove completeness results for WEAK GLOBAL MONOTONICITY and GLOBAL E-MONOTONICITY in the following Sections.

4 Weak Global Monotonicity

In this section, we present a proof for WEAK GLOBAL MONOTONICITY (with explicit representations) along the lines of Park and Darwiche. Note that STRONG GLOBAL MONOTONICITY has been proven to be co-NP^{PP}-complete in [10] using a reduction from the decision variant of the MAP-problem, and that hardness of the weak variant can be proven by restriction. We construct a reduction from A-MAJSAT, the relevant satisfiability variant discussed in Section 2, in order to facilitate our main result in the next section. First, we state the relevant decision problem:

WEAK GLOBAL MONOTONICITY
Instance: Let $\mathbf{B} = (G, \Gamma)$ be a Bayesian network where Γ is composed of explicitly represented rational probabilities, and let Pr be its joint probability distribution. Let $C \in V(G)$ and $E \subseteq V(G) \setminus \{C\}$.
Question: Is $\mathbf{B}$ weakly monotone in distribution in E?

We will see that any instance $(\phi, \mathbf{X_A}, \mathbf{X_M})$ of A-MAJSAT can be translated in a probabilistic network that is monotone, if and only if $(\phi, \mathbf{X_A}, \mathbf{X_M})$ is satisfiable. As an example, let us consider the formula $\phi = \neg(x_1 \wedge x_2) \vee \neg x_3$ (Figure 2), and let $\mathbf{X_A} = \{x_1, x_2\}$ and $\mathbf{X_M} = \{x_3\}$. This is a 'yes'-instance of A-MAJSAT because, for every configuration of $\mathbf{X_A}$, at least half of the configurations of

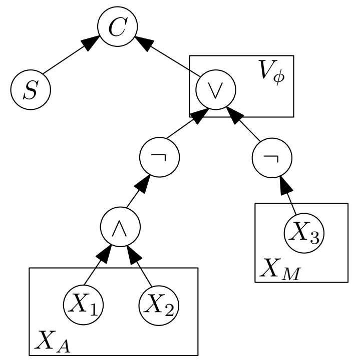

Fig. 2. Construct for hardness proof Monotonicity

$\mathbf{X_M}$ satisfies ϕ. From ϕ we construct a network $\mathbf{B}_\phi$ as described in the previous section. Furthermore, a node C ('classifier') and a node S ('select') is added, with arcs (S, C) and (V_ϕ, C), where V_ϕ is the top node in $\mathbf{B}_\phi$. S has values T and F with uniform distribution, and C has conditional probabilities as denoted in Table 1. We claim, that $\Pr(C \mid S \wedge \mathbf{X_A})$ in the thus constructed network, is weakly monotone in distribution, if and only if the corresponding A-Majsat-instance $(\phi, \mathbf{X_A}, \mathbf{X_M})$ is satisfiable.

	c_1	c_2	c_3
$S = T \wedge V_\phi = T$	0.5	0.25	0.25
$S = T \wedge V_\phi = F$	0.5	0.25	0.25
$S = F \wedge V_\phi = T$	0.25	0.375	0.375
$S = F \wedge V_\phi = F$	0.375	0.5	0.125

Table 1. Conditional probability table for C

Theorem 1. weak Global Monotonicity *is co-*NP$^{\text{PP}}$*-complete*

Proof. To prove membership of co-NP$^{\text{PP}}$, we consider not-weak Global Monotonicity and prove membership of NP$^{\text{PP}}$. In this complement problem we want to know if there exist instantiations to the evidence variables E such that $\mathbf{B}$ is *not* monotone in distribution. This is clearly in NP$^{\text{PP}}$: we can non-deterministically choose instantiations $\mathbf{e_1} \preceq \mathbf{e_2}$ to E and values $c < c' \in \Omega(C)$, and verify that $F_{Pr}(c \mid e_1) \leq F_{Pr}(c' \mid e_1)$, but $F_{Pr}(c' \mid e_2) \leq F_{Pr}(c \mid e_2)$ since Pr is PP-complete.

To prove co-NP$^{\text{PP}}$-hardness, we construct a transformation from A-Majsat. Let $(\phi, \mathbf{X_A}, \mathbf{X_M})$ be an instance of this problem, and let $\mathbf{B}_\phi$ be the network constructed from ϕ as described above. Given a particular configuration $\mathbf{x}$ of all

n variables in $\mathbf{X_A} \cup \mathbf{X_M}$, $\Pr(V_\phi \mid \mathbf{x})$ equals 1 if $\mathbf{x}$ is a satisfying configuration and 0 if it is not, hence, for any configuration $\mathbf{X_A}$, $V_\phi \geq \frac{1}{2}$ if at least half of the instantiations to $\mathbf{X_M}$ satisfy ϕ. Since C is conditioned on V_ϕ, it follows from Table 1 that if any configuration of $\mathbf{X_A}$ leads to $\Pr(V_\phi) < \frac{1}{2}$, then C is no longer monotone in $S \wedge \mathbf{X_A}$, since $F_{Pr}(c_1 \mid S = T) > F_{Pr}(c_1 \mid S = F)$, but $F_{Pr}(c_2 \mid S = T) < F_{Pr}(c_2 \mid S = F)$ as we can calculate[2] from the conditional probability table for C.

Thus, if we can decide whether $\mathbf{B}_\phi$ is weakly globally monotone in $S \cup \mathbf{X_A}$, we are able to decide $(\phi, \mathbf{X_A}, \mathbf{X_M})$. On the other hand, if $(\phi, \mathbf{X_A}, \mathbf{X_M})$ is a satisfying instantiation of A-MAJSAT, then $\Pr(V_\phi) \geq \frac{1}{2}$ and thus $\mathbf{B}_\phi$ is weakly globally monotone. Therefore WEAK GLOBAL MONOTONICITY is co-NP$^{\text{PP}}$-hard. □

5 Global E-Monotonicity

We now use the proof technique from the previous section to prove that GLOBAL E-MONOTONICITY is co-NP$^{\text{NP}^{\text{PP}}}$-complete if we allow implicit representations for the conditional probability distributions, using a reduction from NOT-EA-MAJSAT, which is equivalent to AE-MAJSAT[3]. Again, we start with a formal definition of the relevant decision problem:

GLOBAL E-MONOTONICITY

Instance: Let $\mathbf{B} = (G, \Gamma)$ be a Bayesian network where Γ is composed of rational probabilities, and let Pr be its joint probability distribution. Let $\Omega(V)$ denote the set of values that $V \in V(G)$ can take, and let $C \in V(G)$ and $E \subseteq V(G) \setminus \{C\}$.

Question: Is there an interpretation I_V for all variables $V \in V(G)$, such that $\mathbf{B}$ is monotone in distribution in E?

We will see that any instance $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$ of NOT-EA-MAJSAT can be translated in a probabilistic network for which exists an ordering of the values of its variables that makes the network monotone, if and only if $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$ is *unsatisfiable*. As an example of the GLOBAL E-MONOTONICITY problem, let us consider the formula $\phi = \neg((x_1 \vee x_2) \wedge (x_3 \vee x_4)) \wedge x_5$ (Figure 3), let $\mathbf{X_E} = \{x_1, x_2\}$ and let $\mathbf{X_A} = \{x_3, x_4\}$ and $\overline{\mathbf{X_M}} = \{\neg x_5\}$. One can verify that this is indeed a 'yes'-instance of NOT-EA-MAJSAT: For every configuration of $\mathbf{X_E}$, the configuration $x_3 = x_4 = F$ ensures that at least half of the instantiations of $\overline{\mathbf{X_M}}$ satisfies ϕ. Thus, there does not exist an instantiation to $\mathbf{X_E}$, such that for *all*

[2] $F_{Pr}(c_1 \mid S = T) = \Pr(c_1 \mid V_\phi = T \wedge S = T) \cdot \Pr(V_\phi = T) + \Pr(c_1 \mid V_\phi = F \wedge S = T) \cdot \Pr(V_\phi = F) = (0.5 + \epsilon) \cdot 0.5 + (0.5 - \epsilon) \cdot 0.5 = 0.5$. Likewise, $F_{Pr}(c_1 \mid S = F) = 0.25 \cdot (0.5 - \epsilon) + 0.375 \cdot (0.5 + \epsilon) = 0.3125 + 0.125\epsilon$. On the other hand, $F_{Pr}(c_2 \mid S = T) = \Pr(c_1 \mid S = T) + \Pr(c_2 \mid S = T) = 0.5 + 0.25 < F_{Pr}(c_2 \mid S = F) = \Pr(c_1 \mid S = F) + \Pr(c_2 \mid S = F) = (0.3125 + 0.125\epsilon) + (0.4375 + 0.125\epsilon) = 0.75 + 0.25\epsilon$.

[3] Thus, instead of $\forall^P \exists^P \mathcal{C}$ we use the equivalent problem statement $\neg \exists^P \forall^P \neg \mathcal{C}$. The reader can verify that this is an equivalent problem formulation.

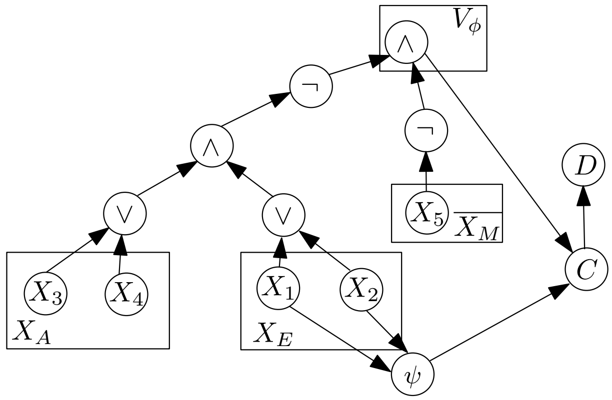

Fig. 3. Construct for hardness proof E-Monotonicity

instantiations to $\mathbf{X_A}$ at least half of the instantiations of $\overline{\mathbf{X_M}}$ does *not* satisfy ϕ.

Again, we denote V_ϕ as the top node in $\mathbf{B}_\phi$. We now add three additional variables, C with values c_1, c_2, c_3, D with values d_1, d_2, and a variable ψ. This variable is implicitly defined and has (implicit) values w_0 to w_{2^m-1} ($m = \mid \mathbf{X_E} \mid$) that correspond to configurations $\mathbf{x_E}$ of $\mathbf{X_E}$. These values are ordered by the binary representation of each configuration $\mathbf{x_E}$, e.g. for an instantiation $\mathbf{x_E} = X_1 = F, \ldots, X_{m-1} = F, X_m = T$ the binary representation would be $0 \ldots 01$ and therefore this particular configuration would correspond with w_1. Likewise, all possible configurations of $\mathbf{X_E}$ are mapped to values w_i of ψ. Furthermore, there are arcs (V_ϕ, C), (ψ, C), (C, D), and from every variable in $\mathbf{X_E}$ to ψ. The conditional probability $\Pr(C \mid V_\phi \wedge \psi)$ is defined in Table 2, where ϵ is a sufficiently small number, e.g. $\epsilon \leq \frac{1}{2^{m+3}}$. The conditional probabilities $\Pr(\psi \mid \mathbf{X_E})$ and $\Pr(D \mid C)$ are defined in Table 3. Note, that the conditional probability distribution of both ψ and C are defined implicitly. The conditional probabilities of D are chosen in such a way, that D is monotone in C if and only if $I_C = \{c_1 < c_2 < c_3\}$. We claim, that there is a possible interpretation I for all variables in $\mathbf{X_E} \cup \{\psi\}$ in the thus constructed network, such that the network is globally monotone, if and only if the corresponding NOT-EA-MAJSAT-instance is satisfiable. To support this claim, we take a closer look at the example. The possible values of ψ are numbered as follows: $w_0 = \{X_1 = F, X_2 = F\}, w_1 = \{X_1 = F, X_2 = T\}, w_2 = \{X_1 = T, X_2 = F\}, w_3 = \{X_1 = T, X_2 = T\}$. For $i = 0 \ldots 3$, the conditional probability table $\Pr(C \mid V_\phi \wedge \psi = w_i)$ is defined as in Table 4. We have already seen that, for all configurations to $\mathbf{X_A}$, the configuration $X_3 = X_4 = F$ of $\mathbf{X_E}$ ensures that the majority of the possible configurations of $\mathbf{X_M}$ satisfies ϕ. Therefore, for all configurations of $\mathbf{X_A}$, there is at least one configuration of ψ (namely, $\psi = w_0$) such that $V_\phi \geq \frac{1}{2}$. Since C is conditioned on V_ϕ, we can calculate from the table that monotonicity is violated: $F_{Pr}(c_1 \mid \psi = w_0) = 0.625 - 0.5\epsilon > F_{Pr}(c_1 \mid \psi = w_1) = 0.4375$ but $F_{Pr}(c_2 \mid \psi = w_0) = 0.75 - 0.5\epsilon < F_{Pr}(c_2 \mid \psi = w_1) = 0.75$.

$$
\begin{array}{rll}
\Pr(C = c_1 \,|\, V_\phi = T \wedge \psi = w_i) = & \frac{1}{2} - \frac{i}{2^{m+1}} - \epsilon & \text{if i = 0} \\
& \frac{1}{2} - \frac{i}{2^{m+1}} & \text{otherwise} \\
\Pr(C = c_2 \,|\, V_\phi = T \wedge \psi = w_i) = & \frac{i+1}{2^m} - \frac{1}{2^{m+1}} & \\
\Pr(C = c_3 \,|\, V_\phi = T \wedge \psi = w_i) = & \frac{1}{2} - \frac{i+1}{2^{m+1}} + \epsilon & \text{if i = 0} \\
& \frac{1}{2} - \frac{i+1}{2^{m+1}} & \text{otherwise} \\
\Pr(C = c_1 \,|\, V_\phi = F \wedge \psi = w_i) = & 1 - \frac{i+1}{2^m} & \\
\Pr(C = c_2 \,|\, V_\phi = F \wedge \psi = w_i) = & \frac{i+1}{2^{m+1}} & \\
\Pr(C = c_3 \,|\, V_\phi = F \wedge \psi = w_i) = & \frac{i+1}{2^{m+1}} &
\end{array}
$$

Table 2. Conditional probability for C

	d_1	d_2
c_1	0.20	0.80
c_2	0.40	0.60
c_3	0.60	0.40

$$
\begin{array}{rll}
\Pr(\psi = w_i \,|\, \mathbf{x_E}) = & 1 & \text{if } w_i \text{ corresponds to } \mathbf{x_E} \\
& 0 & \text{otherwise}
\end{array}
$$

Table 3. Conditional probabilities for $\Pr(D\,|\,C)$ and $\Pr(\psi\,|\,X)$

	c_1	c_2	c_3
$\psi = w_0 \wedge V_\phi = T$	$0.5 - \epsilon$	0.125	$0.375 + \epsilon$
$\psi = w_1 \wedge V_\phi = T$	0.375	0.375	0.25
$\psi = w_2 \wedge V_\phi = T$	0.25	0.625	0.125
$\psi = w_3 \wedge V_\phi = T$	0.125	0.875	0

	c_1	c_2	c_3
$\psi = w_0 \wedge V_\phi = F$	0.75	0.125	0.125
$\psi = w_1 \wedge V_\phi = F$	0.5	0.25	0.25
$\psi = w_2 \wedge V_\phi = F$	0.25	0.375	0.375
$\psi = w_3 \wedge V_\phi = F$	0	0.5	0.5

Table 4. Conditional probability for C in the example

Thus, independent of the way the values of $\Omega(\psi)$ are ordered, there is always at least one violation of monotonicity for any interpretation in $\mathbf{I}_\psi$ if $V_\phi \geq \frac{1}{2}$. If, on the other hand, there does not exist such configuration to $\mathbf{X_E}$, then $V_\phi < \frac{1}{2}$ for all possible configurations of $\mathbf{X_E}$, and thus there is an ordering of the interpretations in $\mathbf{I}_\psi$ such that $\Pr(C\,|\,\mathbf{X_A})$ is monotone. Note that we cannot assume an *a priori* ordering on the values of ψ in this situation: although all configurations of $\mathbf{X_E}$ lead to $V_\phi < \frac{1}{2}$, some may be closer to $\frac{1}{2}$ than others and thus, because of the conditioning on V_ϕ, lead to higher values in C.

Theorem 2. Global E-Monotonicity *is co-*$\text{NP}^{\text{NP}^{\text{PP}}}$*-complete*

Proof. For a membership proof we use not-weak Global Monotonicity as an NP^{PP} oracle. With the aid of this oracle, an interpretation for the values of the variables that violates monotonicity is an NP membership certificate for not-Global E-Monotonicity , thus by definition the problem is in co-$\text{NP}^{\text{NP}^{\text{PP}}}$.

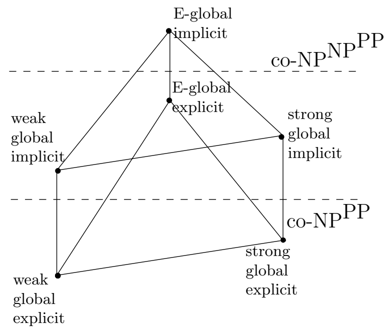

Fig. 4. Known complexity results

To prove co-NP$^{\text{NP}^{\text{PP}}}$-hardness, we construct a transformation from NOT-EA-MAJSAT. Let $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$ be an instance of this problem, and let $\mathbf{B}_\phi$ be the network constructed from ϕ as described above. If $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$ is *not* satisfiable, then there exists an instantiation to ψ, such that $\Pr(V_\psi) \geq \frac{1}{2}$ and thus – again, because of the conditioning of C on V_ψ – monotonicity is violated. But if this is the case, then there exist $w_i, w_j \in \psi$ and $c < c' \in C$ such that $F_{Pr}(c \mid \psi = w_i) \leq F_{Pr}(c' \mid \psi = w_i)$, but $F_{Pr}(c' \mid \psi = w_j) \leq F_{Pr}(c \mid \psi = w_j)$ independent of the ordering of the values of ψ. Note that the variable-and operator-nodes have binary values, making an ordering irrelevant[4], and the ordering on C and D is imposed by the conditional probability $\Pr(D \mid C)$. Thus, if we would be able to decide that there is an interpretation of the values of the variables of $\mathbf{B}_\phi$ such that $\mathbf{B}_\phi$ is globally monotone in distribution, we are able to decide $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$. On the other hand, given that the network is globally monotone, we know that there cannot be an instantiation to $\mathbf{X_E}$ such that $(\phi, \mathbf{X_E}, \mathbf{X_A}, \overline{\mathbf{X_M}})$ is satisfied. Hence, GLOBAL E-MONOTONICITY is co-NP$^{\text{NP}^{\text{PP}}}$-hard.

It may not be obvious that the above construction can be made in polynomial time. Note that, however large X_E may become, both the conditional probabilities $\Pr(\psi \mid \mathbf{X_E})$ and $\Pr(C \mid V_\phi \wedge \psi)$ can be described using only a constant number of bits, since we explicitly allowed Γ to have implicit representations. Therefore, we need only time, polynomial in the size of the input (i.e., the NOT-EA-MAJSAT instance), to construct $\mathbf{B}_\phi$. □

6 Conclusion

In this paper, several variants of the MONOTONICITY problem in probabilistic networks were introduced. In Figure 4, the known complexity results for strong and weak global monotonicity variants, with explicit or implicit conditional probability distribution, and fixed or variable variable orderings are presented. The

[4] if $\mathbf{B}_\phi$ is isotone for $x < \bar{x}$, it is antitone for $\bar{x} < x$ and vice versa

main result is the completeness proof of GLOBAL E-MONOTONICITY with implicit probability representation. It is established that this problem is complete for the class co-$\text{NP}^{\text{NP}^{\text{PP}}}$, a class for which few real-world problems are known to be complete. Unfortunately, a similar complexity result for the variant with explicit representation (or, with a representation where the variables are explicitly defined, while the probabilities are implicit) could not be established. This leaves us with a number of problems that are either in co-NP^{PP}, co-$\text{NP}^{\text{NP}^{\text{PP}}}$, or somewhere in between.

Acknowledgements

The author wishes to thank Hans Bodlaender, Gerard Tel, and Linda van der Gaag, for fruitful discussions on this subject and useful comments on earlier drafts of this paper. Furthermore, the author is very grateful to three anonymous reviewers for their constructive critiques.

References

1. M. R. Garey and D. S. Johnson. *Computers and Intractability. A Guide to the Theory of NP-Completeness*. W. H. Freeman and Co. San Francisco, 1979.
2. B. Kappen, W. Wiegerinck, E. Akay, M. Nijman, J. Neijt, and A. van Beek. Promedas: A diagnostic decision support system. In *Proceedings of the 15th Belgian-Dutch Conference on Artificial Intelligence (BNAIC'03)*, pages 455–456, 2003.
3. M. L. Littman, S. M. Majercik, and T. Pitassi. Stochastic boolean satisfiability. *Journal of Automated Reasoning*, 27(3):251–296, 2001.
4. C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
5. J. D. Park and A. Darwiche. Complexity results and approximation settings for MAP explanations. *Journal of Artificial Intelligence Research*, 21:101–133, 2004.
6. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, Palo Alto, 1988.
7. D. Roth. On the hardness of approximate reasoning. *Artificial Intelligence*, 82(1-2):273–302, 1996.
8. S. E. Shimony. Finding MAPs for belief networks is NP-hard. *Artificial Intelligence*, 68(2):399–410, 1994.
9. J. Torán. Complexity classes defined by counting quantifiers. *Journal of the ACM*, 38(3):752–773, 1991.
10. L. C. van der Gaag, H. L. Bodlaender, and A. Feelders. Monotonicity in Bayesian networks. In *Twentieth Conference on Uncertainty in Artificial Intelligence*, pages 569–576. AUAI Press, 2004.
11. L. C. van der Gaag, S. Renooij, C. L. M. Witteman, B. M. P. Aleman, and B. G. Taa. Probabilities for a probabilistic network: a case study in oesophageal cancer. *Artificial Intelligence in Medicine*, 25:123–148, 2002.
12. K. W. Wagner. The complexity of combinatorial problems with succinct input representation. *Acta Informatica*, 23:325–356, 1986.