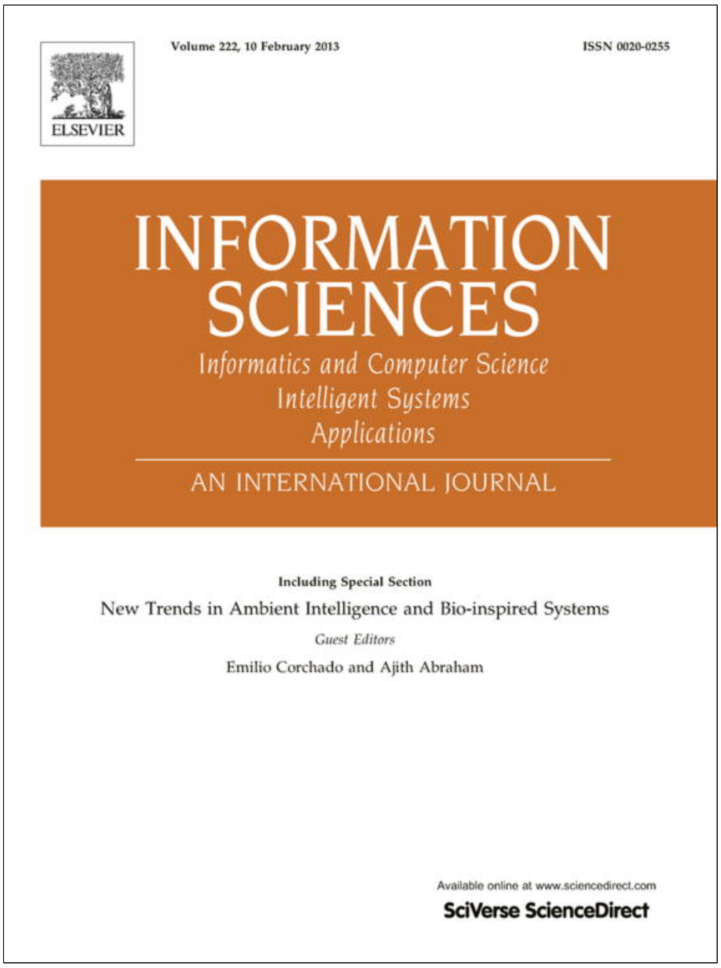
ELSEVIER
Volume 222, 10 February 2013
ISSN 0020-0255
INFORMATION SCIENCES
Informatics and Computer Science
Intelligent Systems
Applications
AN INTERNATIONAL JOURNAL
Including Special Section
New Trends in Ambient Intelligence and Bio-inspired Systems
Guest Editors
Emilio Corchado and Ajith Abraham
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Information Sciences 222 (2013) 439–458

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

The effect of block parameter perturbations in Gaussian Bayesian networks: Sensitivity and robustness

Miguel A. Gómez-Villegas [a], Paloma Main [a,*], Rosario Susi [b]

[a] Departamento de Estadística e Investigación Operativa, Universidad Complutense de Madrid, 28040 Madrid, Spain
[b] Departamento de Estadística e Investigación Operativa III, Universidad Complutense de Madrid, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 14 December 2010
Received in revised form 2 July 2012
Accepted 3 August 2012
Available online 17 August 2012

Keywords:
Decision support system
Gaussian Bayesian network
Sensitivity analysis
Robustness analysis

ABSTRACT

In this work we study the effects of model inaccuracies on the description of a Gaussian Bayesian network with a set of variables of interest and a set of evidential variables. Using the Kullback–Leibler divergence measure, we compare the output of two different networks after evidence propagation: the original network, and a network with perturbations representing uncertainties in the quantitative parameters. We describe two methods for analyzing the sensitivity and robustness of a Gaussian Bayesian network on this basis.

In the sensitivity analysis, different expressions are obtained depending on which set of parameters is considered inaccurate. This fact makes it possible to determine the set of parameters that most strongly disturbs the network output. If all of the divergences are small, we can conclude that the network output is insensitive to the proposed perturbations. The robustness analysis is similar, but considers all potential uncertainties jointly. It thus yields only one divergence, which can be used to confirm the overall sensitivity of the network. Some practical examples of this method are provided, including a complex, real-world problem.

1. Introduction

A Bayesian network (BN) is a probabilistic graphical model used to study a set of variables with a known dependence structure. Probabilistic graphical models are framed in the field of AI to model uncertainty and numerous techniques have been developed to improve models and its learning. Learning an optimal Bayesian network classifier is an NP-hard problem then some algorithms has been developed to propose improvements to the traditional models (see [17]). Schemes to refine rules and methods to develop learning algorithms reducing computational cost and improving learning accuracy can be found in [31,32]. Computational intelligence techniques, as neural networks, support vector machines and extreme learning machine (see [15]), have been used in many applications. Our interest models, BNs, are revealed to be one of the best technique to study this type of real world problems in a number of complex domains, including medical diagnosis or dynamic systems, for example.

A BN has two components, as shown in Definition 1. The qualitative part is a directed acyclic graph (DAG) showing the dependence structure of the variables. The quantitative part consists of conditional probability distributions assigned to the problem variables giving their parents in the DAG.

* Corresponding author.
E-mail address: pmain@mat.ucm.es (P. Main).

0020-0255/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2012.08.004

Building a BN is a complex task because it requires the specification of a large number of parameters subject to cognitive biases [2]. The parameters are generally estimated from statistical data or assessed by human experts in the domain of application. ‘‘Experts are often reluctant to assess all the parameters required, feeling that they are unable to give assessments with a high level of accuracy’’ [6]. As a consequence of incompleteness of data and partial knowledge of the domain, the assessments obtained inevitably are inaccurate [30]. With inaccurate parameters, network output after evidence propagation may also be inaccurate, depending on the sensitivity of the model. Then, sensitivity analyses in BNs are necessary to evaluate the effect of uncertainty in the network and to determine the values of the parameters to get accurate network outputs.

Sensitivity analysis is a general technique for investigating the robustness of the output of a mathematical model and is performed for various different purposes. The practicability of conducting such an analysis of a probabilistic network has recently been studied extensively, resulting in a variety of new insights and effective methods. A survey with some of these research results is in van der Gaag et al. [30].

Authors like Laskey [21], Coupé et al. [6,7], Chan and Darwiche [5] or Bednarski et al. [1] have studied the sensitivity in discrete BNs, where the parameters are the conditional probabilities of variables given its parents in the DAG. Moreover, all the papers mentioned above deal with variations in one parameter at a time holding the other fixed and with only one interest variable, being the network output the probability distribution of the interest variable given the evidence.

An analysis considering the effects of variations in values of only one parameter, holding the other fixed, is called *one way sensitivity analysis*. Then, next sensitivity analyses are one way sensitivity analyses.

Laskey [21] presents a methodology for analytic computation of sensitivity values in discrete BNs when there is only one variable of interest in the problem. These sensitivity values measure the impact of small changes in the parameter on the probability value of interest computing partial derivatives of output probabilities with respect to inaccurate parameters.

Coupé et al. [6] provide an elicitation procedure to assess values to the parameters in which, alternatively, sensitivity analyses are performed and parameters assessments redefined. Given the highly time-consuming in [7] parameters that cannot affect the output are studied.

Moreover, for discrete BNs, where parameters are conditional probabilities, [5] propose a distance measure between two probability distributions to compute the amount of change that occurs when moving from one distribution to another. They contrast the proposed measure with classical measures as the Kullback–Leibler (KL) divergence [20] and show that belief change between two states of belief can be unbounded, even when their KL divergence tends to zero. They show, however, that KL divergence can be used to compute the average change in beliefs. Despite this inappropriate behavior in discrete BN it is shown [9] that KL measure is a good one to describe the effects of parameter perturbations in Gaussian models.

Finally, Bednarski et al. [1] focus their one way sensitivity analysis in identifying the set of sensibilities that affects the variable of interest.

When the interest is about a set of parameters, then, the objective is to analyze the effects of variations in values of a set of parameters at the same time, this analysis is called *n-way sensitivity analysis*. Authors like KJærulff and van der Gaag [19] or Chan et al. [4] introduce *n*-way sensitivity analyses to identify multiple parameters changes in discrete BNs. In [4] only one interest variable is considered while authors in [19] work with a set of interest variables.

Literature about sensitivity analysis in Gaussian Bayesian networks (GBNs) is not extensive. Authors like Castillo and KJærulff [3] or Gómez-Villegas et al. [9,10] have studied the problem of uncertainty in parameters assignments in GBNs.

In [3] a one way sensitivity analysis based on [21] is proposed. Then the impact of small changes in the network parameters is studied evaluating local aspects of the distribution such as location and dispersion. In this analysis only one variable of interest is considered. Moreover, the analysis also focuses on small changes about the parameters of the network and evaluates the impact of uncertainty on mean vector and covariance matrix.

In [9] a one way sensitivity analysis is proposed to evaluate the effects of small and large changes in the network parameters considering a global sensitivity measure. As in [5] this is a distance measure between two probability distributions, but in this context – GBNs – the measure used is the KL divergence.

Moreover, in [10], the expressions obtained for the sensitivity analysis were evaluated in the limit, considering extreme changes in the network parameters.

All the papers mentioned above deal with variations in one parameter at a time holding the others fixed. Then, both analyses are one way sensitivity analysis.

The present paper aims to generalize the sensitivity analysis presented in [9] in two ways. First, by developing an *n*-way version of the sensitivity analysis we hope to study the effects of perturbations in a set of parameters. Second, considering a GBN with several variables of interest and evidential variables.

When analyzing the sensitivity of the network, we normally select a set of parameters to be reviewed. But it could also be necessary to determine the simultaneous impact of all inaccuracies over the whole network. With this aim, we propose a robustness analysis. This paper therefore offers two analysis methods: one for sensitivity, and another for robustness.

Other works have focused on uncertainty about the arcs of the DAG that describes the network. Authors like Renooij [28] for discrete BNs or Gómez-Villegas et al. [11] for GBNs develop analyses to study the effect of adding or removing an arc of the DAG, changing the dependence structure of the network. In both cases, the KL divergence is used to compare probability distributions.

Thus, in a general framework we compute the KL divergence of the network output after evidence propagation under two different models, the original and the perturbed, to evaluate the effect of inaccuracies in the assigned parameters.

If all the divergences are small we can conclude that the network is not sensitive to the proposed perturbations. A similar methodology for the robustness analysis is developed, but working with only one perturbed model where all inaccuracies are considered at the same time.

Both analyses are applied to a GBN with various types of uncertainty. A complex, practical problem dealing with structure reliability of a building is also studied.

This paper is organized as follows. In Section 2 we introduce BNs and GBNs. In Section 3, we describe our proposed methodology for evaluating the sensitivity and robustness of a GBN. In Section 4 we present the results of these methods. In Section 5 we introduce a simple GBN, and study its sensitivity and robustness under two different cases of uncertainty. In Section 6, we apply the proposed methodology to a complex network. Finally, Section 7 summarizes our conclusions and proofs with details about calculations of the proposed results are given in Appendix A.

2. GBNs and evidence propagation

Throughout this paper, random variables will be denoted by capital letters and their values by lowercase letters. In the multidimensional case, boldface characters will be used.

Definition 1 (*Bayesian network (BN)*). A BN is a couple $(\mathcal{G}, \mathcal{P})$, where $\mathcal{G}$ is a DAG in which the nodes represent ordered random variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ and the edges represent probabilistic dependencies and $\mathcal{P} = \{P(X_1|pa(X_1)), \ldots, P(X_n|pa(X_n))\}$ is a set of conditional probability distributions, where $pa(X_i)$ is the set of parents of node X_i in $\mathcal{G}$, $pa(X_i) \subseteq \{X_1, \ldots, X_{i-1}\}$.

The set $\mathcal{P}$ defines the associated joint probability distribution:

$$P(\mathbf{X}) = \prod_{i=1}^{n} P(X_i|pa(X_i)). \tag{1}$$

Among others, BNs have been studied by Pearl [26], Lauritzen [22], Cowell et al. [8] and Jensen et al. [16].

Definition 2 (*Gaussian Bayesian network (GBN)*). A GBN is a BN where the joint probability density associated with the variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ is a multivariate normal distribution $N(\mu, \boldsymbol{\Sigma})$ given by

$$f(\mathbf{x}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mu)\right\}.$$

$\mu$is the n-dimensional mean vector and $\boldsymbol{\Sigma}$ is the $n \times n$ positive definite covariance matrix. The conditional probability density for X_i $(i = 1 \ldots, n)$ that satisfies (1) is the univariate normal distribution given by

$$X_i|pa(X_i) \sim N\left(\mu_i + \sum_{j=1}^{i-1}\beta_{ji}(x_j - \mu_j), v_i\right) \tag{2}$$

where μ_i is the mean of X_i, β_{ji} is the regression coefficient of X_j when X_i is regressed on its parents, and v_i is the conditional variance of X_i given its parents.

It can be pointed out that if $\beta_{ji} = 0$ then there is no link from X_j to X_i. To get the covariance matrix $\boldsymbol{\Sigma}$ from $\{v_i\}$ and $\{\beta_{ji}\}$ we refer to the algorithm proposed by Shacker and Kenley [29] where the matrices **D** and **B** are defined next. Let **D** be a diagonal matrix with the conditional variances v_i, $\mathbf{D} = diag(v)$. Let **B** be a strictly upper triangular matrix with the regression coefficients β_{ji} where X_j is a parent of X_i, for the variables in $\mathbf{X}$with $j < i$. Then, the covariance matrix $\boldsymbol{\Sigma}$ can be computed as

$$\boldsymbol{\Sigma} = [(\mathbf{I}-\mathbf{B})^{-1}]^T\mathbf{D}(\mathbf{I}-\mathbf{B})^{-1} \tag{3}$$

For the calculations, the open source programming language and environment for statistical computing and graphics R [27], is used.

Example 1. This problem models the amount of time that a machine will function before failing. The machine is made up of 5 elements, connected as shown in the DAG of Fig. 1, where the numbers of the edges are the coefficients β_{ji}. For example, the time before failing of X_3 depends on the time to fail of both, X_1 and X_2, with coefficients 2 and 1 respectively. Their joint working time is assumed to be normally distributed, as is the working time of each element. Experts are interested in predicting how long elements 3, 4 and 5 will continue working before they fail. Therefore, our main interest is in variables X_3, X_4 and X_5.

The quantitative part of the network, given by the conditional distributions in (2), is defined by

$$\begin{aligned}
&X_1 \sim N(2,3)\\
&X_2 \sim N(3,2)\\
&X_3|X_1 = x_1, X_2 = x_2 \sim N(3 + 2(x_1 - 2) + 1(x_2 - 3), 1)\\
&X_4 \sim N(4,2)\\
&X_5|X_3 = x_3, X_4 = x_4 \sim N(5 + 1(x_3 - 3) + 2(x_4 - 4), 3)
\end{aligned}$$

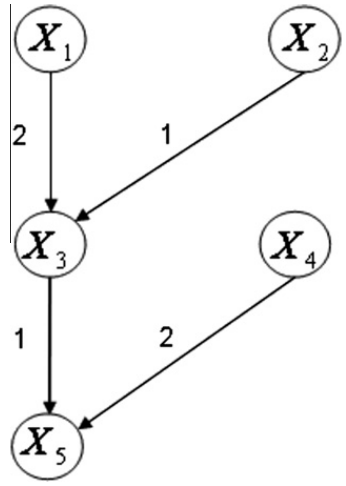

Fig. 1. DAG of Example 1.

The joint probability distribution of **X** = {X_1,X_2,X_3,X_4,X_5} is therefore given by a multivariate normal distribution **X** $\sim N(\mu, \mathbf{\Sigma})$, with

$$\mu = \begin{pmatrix} 2 \\ 3 \\ 3 \\ 4 \\ 5 \end{pmatrix}; \quad \mathbf{\Sigma} = \begin{pmatrix} 3 & 0 & 6 & 0 & 6 \\ 0 & 2 & 2 & 0 & 2 \\ 6 & 2 & 15 & 0 & 15 \\ 0 & 0 & 0 & 2 & 4 \\ 6 & 2 & 15 & 4 & 26 \end{pmatrix}$$

where the elements of the mean vector $\mu$are μ_i for all X_i and the matrix has been calculated using (3).

As can be seen, all the connected variables have a nonzero value in the covariance matrix $\mathbf{\Sigma}$. In addition to the arcs that appear in the DAG that collect relations between parents and children, there are connections in serie between the variables X_1 and X_5 through X_3 as well as between X_2 and X_5. Also, two convergent connections exist because X_1 and X_2 have a common effect, X_3, as well as X_3 and X_4 have a common child, X_5. Consequently, X_1 and X_5 are independent given X_3 and so are X_2 and X_5. However, X_1 and X_2 are conditionally dependent given X_3 as well as X_3 and X_4 are also conditionally dependent given X_5. More information about connections in DAGs can be found in [16].

2.1. Evidence propagation in GBNs

The main result associated with BNs is the inference process known as evidence propagation. If there exists evidence or information about the value of a variable in the problem, this mechanism updates the probability distribution of other variables in the network.

Several algorithms have been developed to propagate evidence in BNs. For GBNs, some of these are based on methods developed for discrete BNs. Two examples are the algorithm introduced by Xu and Pearl [34], and the propagation method proposed by Normand and Tritchler [25] using Pearl's polytree algorithm [26].

Otherwise, evidence propagation in GBNs is generally based on computing the conditional probability distribution of a multivariate normal distribution given a set of evidential variables, i.e., given a set of variables with known states. Then, we can consider **X** = (**Y**, **E**)′, where **Y** is the set of variables of interest and **E** is the set of evidential variables. **e** is the evidence about the variables in **E**.

For simple and efficient propagation, an incremental method is used to update one evidential node at a time [3]. After the evidence propagation the network output is given by the distribution of **Y**|**E**, being **Y**|**E** $\sim N(\mu^{\mathbf{Y}|\mathbf{E}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$ with parameters:

$$\begin{aligned} \mu^{\mathbf{Y}|\mathbf{E}} &= \mu_{\mathbf{Y}} + \mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1}(\mathbf{e} - \mu_{\mathbf{E}}) \\ \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}} &= \mathbf{\Sigma}_{\mathbf{YY}} - \mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1}\mathbf{\Sigma}_{\mathbf{EY}} \end{aligned} \tag{4}$$

A particular situation is given in the following example.

Example 2. Working with the GBN introduced in Example 1, presume that we have evidence about the values of X_1 and X_2. Let **E** = {X_1 = 2, X_2 = 4} represent this evidence.

The incremental method considers $X_1 = 2$, computes the distribution for X_2, X_3, X_4, $X_5|X_1 = 2$, and only then considers $X_2 = 4$. After evidence propagation, the probability distribution of the variables of interest $\mathbf{Y} = (X_3, X_4, X_5)$, given the evidence $\mathbf{E} = \{X_1 = 2, X_2 = 4\}$, i.e. the network output, is $\mathbf{Y}|\mathbf{E} \sim N(\mu^{\mathbf{Y}|\mathbf{E}}, \boldsymbol{\Sigma}^{\mathbf{Y}|\mathbf{E}})$ with parameters

$$\mu^{\mathbf{Y}|\mathbf{E}} = \begin{pmatrix} 4 \\ 4 \\ 6 \end{pmatrix}; \quad \boldsymbol{\Sigma}^{\mathbf{Y}|\mathbf{E}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 1 & 4 & 12 \end{pmatrix}$$

3. Sensitivity and robustness

The methods developed in this work essentially consist of comparing the network output under two different models, the original model and the perturbed model.

The original model is given by the initial parameters associated with the quantitative part of the GBN, i.e., the initial values estimated from statistical data or assigned by experts to μ and $\boldsymbol{\Sigma}$. The perturbed model quantifies uncertainty in the parameters by introducing two sets of additive perturbations: δ for the mean, and $\boldsymbol{\Delta}$ for the covariance matrix.

Thus, to achieve an n-way sensitivity analysis, we must compare five different types of perturbed models to the original model. The type of model depends on which parameters are thought to be inaccurate. This construction generalizes the result reported by [9]. To obtain the five perturbed models, consider the following partition of the perturbations

$$\delta = \begin{pmatrix} \delta_{\mathbf{Y}} \\ \delta_{\mathbf{E}} \end{pmatrix} \quad \boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{\mathbf{YY}} & \boldsymbol{\Delta}_{\mathbf{YE}} \\ \boldsymbol{\Delta}_{\mathbf{EY}} & \boldsymbol{\Delta}_{\mathbf{EE}} \end{pmatrix}$$

- When there is uncertainty in the mean of the variables of interest, the perturbed model is $\mathbf{X} \sim N(\mu^{\delta_{\mathbf{Y}}}, \boldsymbol{\Sigma})$ with

$$\mu^{\delta_{\mathbf{Y}}} = \begin{pmatrix} \mu_{\mathbf{Y}} + \delta_{\mathbf{Y}} \\ \mu_{\mathbf{E}} \end{pmatrix}$$

- When there is uncertainty in the mean of the evidential variables, the perturbed model is $\mathbf{X} \sim N(\mu^{\delta_{\mathbf{E}}}, \boldsymbol{\Sigma})$ being

$$\mu^{\delta_{\mathbf{E}}} = \begin{pmatrix} \mu_{\mathbf{Y}} \\ \mu_{\mathbf{E}} + \delta_{\mathbf{E}} \end{pmatrix}$$

- When there is uncertainty in the variances and covariances between the variables of interest, the perturbed model is $\mathbf{X} \sim N(\mu, \boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YY}}})$. The covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YY}}}$ must be positive definite, and is given by

$$\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YY}}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{YY}} + \boldsymbol{\Delta}_{\mathbf{YY}} & \boldsymbol{\Sigma}_{\mathbf{YE}} \\ \boldsymbol{\Sigma}_{\mathbf{EY}} & \boldsymbol{\Sigma}_{\mathbf{EE}} \end{pmatrix}$$

- When there is uncertainty in the variances and covariances between the evidential variables, the perturbed model is $\mathbf{X} \sim N(\mu, \boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{EE}}})$. The covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{EE}}}$ must be positive definite, and has the following form:

$$\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{EE}}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{YY}} & \boldsymbol{\Sigma}_{\mathbf{YE}} \\ \boldsymbol{\Sigma}_{\mathbf{EY}} & \boldsymbol{\Sigma}_{\mathbf{EE}} + \boldsymbol{\Delta}_{\mathbf{EE}} \end{pmatrix}$$

- When there is uncertainty in the covariances between the variables of interest and the evidential variables, the perturbed model is $\mathbf{X} \sim N(\mu, \boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YE}}})$. The covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YE}}}$ must be positive definite, and has the following form:

$$\boldsymbol{\Sigma}^{\boldsymbol{\Delta}_{\mathbf{YE}}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{YY}} & \boldsymbol{\Sigma}_{\mathbf{YE}} + \boldsymbol{\Delta}_{\mathbf{YE}} \\ \boldsymbol{\Sigma}_{\mathbf{EY}} + \boldsymbol{\Delta}_{\mathbf{YE}}^T & \boldsymbol{\Sigma}_{\mathbf{EE}} \end{pmatrix}$$

The KL divergence is used to compare the network output of the perturbed model, obtained as shown before, with the network output of the original model. Obviously outputs of the networks are obtained after evidence propagation. We thereby obtain five different divergence measures, one for each of the perturbed models considered.

If all the divergences are close to zero, it can be concluded that the network is not sensitive to any of the proposed perturbations. In this case, the network is said to be robust to the proposed perturbations. Of course, even in this case the divergences could be useful to compare different networks or sensitivity to different sets of perturbations in the same network.

This paper also develops a method for analyzing the global robustness that considers only one perturbed model. This analysis evaluates the effects of uncertainty and inaccuracy in all parameters simultaneously. Specifically, the perturbed model is obtained by adding all perturbations, $\mathbf{X} \sim N(\mu^{\delta}, \boldsymbol{\Sigma}^{\boldsymbol{\Delta}}) = N(\mu + \delta, \boldsymbol{\Sigma} + \boldsymbol{\Delta})$, where δ is a perturbation to the mean vector and $\boldsymbol{\Delta}$ is a perturbation to the covariance matrix. Evidence propagation is then performed and their outputs are compared using the KL divergence, as with the sensitivity analysis. This divergence provides a global measure of the difference between two probability distributions, rather than comparing their local features, and it is useful for comparisons with other uncertain situations.

The KL divergence evaluates the amount of information available to discriminate between two probability distributions, and it is used to compare the overall behavior of the considered probability distributions. Furthermore, this measure is useful when researcher has no idea about which characteristics of the variables of interest are most representative.

4. Main results

The original network output is given by the conditional probability density f of the distribution $\mathbf{Y}|\mathbf{E} \sim N(\mu^{\mathbf{Y}|\mathbf{E}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$. The perturbed network output is given by the conditional probability density $f^{\mathbf{p}_j}$ of the distribution $\mathbf{Y}|\mathbf{E}, p_j \sim N(\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})$. Comparing the two with the KL divergence, the expectation with respect to the density f is as follows,

$$KL^{\mathbf{p}_j}(f^{\mathbf{p}_j}|f) = E_f\left[\ln\frac{f}{f^{\mathbf{p}_j}}\right] = \frac{1}{2}\left[-\ln|\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}| + trace\left(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}\right) - \dim(\mathbf{Y})\right] + \frac{1}{2}\left[(\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j} - \mu^{\mathbf{Y}|\mathbf{E}})^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}(\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j} - \mu^{\mathbf{Y}|\mathbf{E}})\right] \quad (5)$$

For calculations as well as some other interesting aspects of this divergence measure for Gaussian distributions see [33].

Also it can be pointed that direct evaluation of KL divergence for Gaussian distributions has a computational complexity of order $O(q^3)$ being q the dimension of $\mathbf{Y}$, the vector of interest variables. Therefore, the difference in usage between single perturbation variable and a set of perturbation variables depends on the size of the set of variables of interest because that is the dimension of the square matrices involved in calculations.

The next subsections give a detailed analysis of the different types of perturbations with some comparisons of the output effects they can produce. For the proofs see Appendix A.

4.1. Uncertainty about the mean vector

In Proposition 1 we detail the KL divergence when uncertainty is displayed only in the mean vector and may be in the mean of $\mathbf{Y}$ or the mean of $\mathbf{E}$. Then, in Proposition 2 we introduce an upper limit for such measures, so that will limit the values of KL divergence for any value delta perturbation.

Proposition 1. *Let $(\mathcal{G}, \mathcal{P})$ be a GBN with $\mathbf{X} \sim N(\mu, \mathbf{\Sigma})$, where $\mathbf{X} = \{\mathbf{Y}, \mathbf{E}\}$, $\mathbf{Y}$being the set of variables of interest and $\mathbf{E}$ being the set of evidential variables. For δ, the perturbation of the mean vector $\delta = (\delta_{\mathbf{Y}}, \delta_{\mathbf{E}})^T$, the following results hold.*

1. *When the perturbation $\delta_{\mathbf{Y}}$ is added to the mean vector of the variables of interest, after evidence propagation the perturbed model is $\mathbf{Y}|\mathbf{E}, \delta_Y \sim N(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{Y}}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$, with $\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{Y}}} = \mu^{\mathbf{Y}|\mathbf{E}} + \delta_{\mathbf{Y}}$, and the KL divergence is given by*

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) = \frac{1}{2}\left[\delta_{\mathbf{Y}}^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}\delta_{\mathbf{Y}}\right]$$

2. *When the perturbation $\delta_{\mathbf{E}}$ is added to the mean vector of the evidential variables, the perturbed model after evidence propagation is $\mathbf{Y}|\mathbf{E}, \delta_E \sim N(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{E}}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$, with $\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{E}}} = \mu^{\mathbf{Y}|\mathbf{E}} - \mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1}\delta_{\mathbf{E}}$, and the KL divergence is*

$$KL^{\mu_{\mathbf{E}}}(f^{\mu_{\mathbf{E}}}|f) = \frac{1}{2}\left[\delta_{\mathbf{E}}^T(\mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1})^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}(\mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1})\delta_{\mathbf{E}}\right]$$

When the uncertainty is only about the mean vector, the expression in (5) is reduced to the mean vector components. Therefore both expressions in Proposition 1 are similar in the sense that the KL divergence is half the product of the transposed perturbation by some components of the covariance matrix by the perturbation ($\boldsymbol{\delta}_Y$ or $\boldsymbol{\delta}_E$). In both cases the inverse of the covariance matrix appears after evidence propagation. To fully understand the expressions described in Proposition 1, a single perturbation can be considered in the vector $\boldsymbol{\delta}_Y$ or $\boldsymbol{\delta}_E$, so that vector consists of zeros except a value nonzero. In that case, it is observed that the perturbation appears in quadratic form, being the KL divergence a parabola centered at zero (when there is no perturbation). For a more detailed analysis of divergences when uncertainty appears in the mean of a single node in the network, you can see [10].

The KL divergence obtained for the inaccurate mean vector of the evidential variables is the same as the divergence measure for a perturbation of the evidence vector **e**. This gives us a tool to evaluate the influence of extreme evidence on the network output.

To establish some bounds on these divergence measures we can use the Cauchy–Schwarz inequality that gives rise to the Maximization Lemma for positive definite matrices [18].

Proposition 2. *For perturbations in the mean vector next maximum values for the divergence measure can be obtained*

1. *If the mean vector of the variables of interest is perturbed with some perturbations $\delta_{\mathbf{Y}}$ such that its Euclidean norm $\|\delta_{\mathbf{Y}}\|_2^2 = \delta_{\mathbf{Y}}^T\delta_{\mathbf{Y}} \leqslant r^2$ it can be shown*

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) \leqslant \frac{\|\delta_{\mathbf{Y}}\|_2^2}{2\lambda_q} \leqslant \frac{r^2}{2\lambda_q} \quad (6)$$

if $\lambda_1 \geqslant \cdots \geqslant \lambda_q$ are the ordered eigenvalues of $\Sigma^{Y|E}$. Also the maximum is attained with $r * e_q$ being e_q the corresponding eigenvector.

2. *For the case of perturbations δ_E in the mean vector of evidence variables such that,*

$$\left\|\Sigma_{YE}\Sigma_{EE}^{-1}\delta_E\right\|_2^2 = \|B_{Y|E}\delta_E\|_2^2 \leqslant r^{*2}$$

it holds

$$KL^{\mu_E}(f^{\mu_E}|f) \leqslant \frac{\|B_{Y|E}\delta_E\|_2^2}{2\lambda_q} \leqslant \frac{r^{*2}}{2\lambda_q} \tag{7}$$

*if $\lambda_1 \geqslant \cdots \geqslant \lambda_q$ are the ordered eigenvalues of $\Sigma^{Y|E}$. Also it can be pointed that $B_{Y|E} = \Sigma_{YE}\Sigma_{EE}^{-1}$ is the matrix of regression coefficients of **Y** on **E**.*

The expressions introduced in Proposition 2 lead us to calculate an upper bound for the KL divergence. These results are useful because they may set an upper bound to the divergence given the uncertainty.

4.2. Uncertainty about the covariance matrix

In Propositions 3 and 4 we studied uncertainty, with the KL divergence, when the inaccurate parameters are in the covariance matrix. For uncertainties related to the covariance matrix, the covariance matrices of the perturbed models have to be positive definite. Moreover, Proposition 4 leads us to fix an upper limit for expressions obtained in Proposition 3, so that will limit the values of KL divergence for some perturbations associated to uncertainty of the covariance matrix.

Proposition 3. *Let (G,P) be a GBN with $X \sim N(\mu, \Sigma)$ and **X** = {**Y**, **E**}, **Y** being the set of variables of interest and **E** being the set of evidential variables. The perturbed covariance matrix Δ is*

$$\Delta = \begin{pmatrix} \Delta_{YY} & \Delta_{YE} \\ \Delta_{EY} & \Delta_{EE} \end{pmatrix}$$

All the covariance matrices considered in this framework have to be positive definite. The results are as follows:

1. *When a perturbation Δ_{YY} is added to the variances and covariances of the variables of interest, the perturbed model after evidence propagation is given by $Y|E, \Delta_{YY} \sim N(\mu^{Y|E}, \Sigma^{Y|E,\Delta_{YY}})$, with $\Sigma^{Y|E,\Delta_{YY}} = \Sigma^{Y|E} + \Delta_{YY}$, and the KL divergence is given by*

$$KL^{\Sigma_{YY}}(f^{\Sigma_{YY}}|f) = \frac{1}{2}\left[-\ln|\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{YY}})^{-1}| + trace(\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{YY}})^{-1}) - \dim(Y)\right]$$

2. *When a perturbation Δ_{EE} is added to the variances and covariances of the evidential variables, the perturbed model after evidence propagation is given by $Y|E, \Delta_{EE} \sim N(\mu^{Y|E,\Delta_{EE}}, \Sigma^{Y|E,\Delta_{EE}})$, with $\mu^{Y|E,\Delta_{EE}} = \mu_Y + \Sigma_{YE}(\Sigma_{EE} + \Delta_{EE})^{-1}(e - \mu_E)$ and $\Sigma^{Y|E,\Delta_{EE}} = \Sigma_{YY} - \Sigma_{YE}(\Sigma_{EE} + \Delta_{EE})^{-1}\Sigma_{EY}$. The KL divergence is*

$$KL^{\Sigma_{EE}}(f^{\Sigma_{EE}}|f) = \frac{1}{2}\left[-\ln|\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{EE}})^{-1}| + trace(\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{EE}})^{-1}) - \dim(Y)\right] + \frac{1}{2}[(\mu^{Y|E,\Delta_{EE}} - \mu^{Y|E})^T(\Sigma^{Y|E,\Delta_{EE}})^{-1}(\mu^{Y|E,\Delta_{EE}} - \mu^{Y|E})]$$

3. *When a perturbation Δ_{YE} is added to the covariances between **Y** and **E**, the perturbed model af ter evidence propagation is $Y|E, \Delta_{YE} \sim N(\mu^{Y|E,\Delta_{YE}}, \Sigma^{Y|E,\Delta_{YE}})$, with $\mu^{Y|E,\Delta_{YE}} = \mu_Y + (\Sigma_{YE} + \Delta_{YE})\Sigma_{EE}^{-1}(e - \mu_E)$ and $\Sigma^{Y|E,\Delta_{YE}} = \Sigma_{YY} - (\Sigma_{YE} + \Delta_{YE})\ \Sigma_{EE}^{-1}(\Sigma_{EY} + \Delta_{EY})$. The KL divergence is*

$$KL^{\Sigma_{YE}}(f^{\Sigma_{YE}}|f) = \frac{1}{2}\left[-\ln|\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{YE}})^{-1}| + trace(\Sigma^{Y|E}(\Sigma^{Y|E,\Delta_{YE}})^{-1}) - \dim(Y)\right] + \frac{1}{2}\left[(e - \mu_E)^T(\Sigma_{EE}^{-1})^T\Delta_{YE}^T(\Sigma^{Y|E,\Delta_{YE}})^{-1}\Delta_{YE}\Sigma_{EE}^{-1}(e - \mu_E)\right].$$

As can be seen in Proposition 3, when the uncertainty is between the variances and covariances of **Y**, i.e., Δ_{YY} is nonzero, the covariance matrix after the evidence propagation is the only element to be disturbed. However, for the remaining cases, perturbation is reflected both in the mean vector and in the covariance matrix after the evidence propagation.

Therefore, the first expression of Proposition 3 corresponds only to the first part of the expression in (5), depending on the logarithm of the product of the inverse of the covariance matrix given the uncertainty by the matrix covariance without uncertainty, in both cases after the evidence propagation and depending on the trace of this product. Moreover, this expression depends on the dimension of **Y**.

In all other expressions of Proposition 3, corresponding to uncertainty covariance of **E**, then $\Delta_{EE} \neq 0$, and between covariance of **Y** and **E**, where $\Delta_{YE} \neq 0$, in addition to the terms discussed above, it also shows the average component of (5).

To establish some bounds for the cases of covariance submatrices perturbations some concepts have to be fixed because we have to obtain positive definite covariance matrices. Firstly, for additive perturbations the notion of interval matrix can be used [12]. Let $\mathbf{\Sigma}^*$ and $\mathbf{\Delta}$ be symmetric real $n \times n$ matrices, $\mathbf{\Delta} \geqslant \mathbf{0}$. The set of matrices

$$\mathbf{\Sigma}^I = [\mathbf{\Sigma}^* - \mathbf{\Delta}, \mathbf{\Sigma}^* + \mathbf{\Delta}] = \{\mathbf{\Sigma}; \mathbf{\Sigma}^* - \mathbf{\Delta} \leqslant \mathbf{\Sigma} \leqslant \mathbf{\Sigma}^* + \mathbf{\Delta}\}$$

where the inequalities are to be understood componentwise, is called a symmetric interval matrix. As a first step, conditions of positive definiteness of interval matrices [13,14] are given by

$$\begin{aligned} &\rho(|\mathbf{\Sigma}^{*-1}|\mathbf{\Delta}) < 1 \text{ or} \\ &\rho(\mathbf{\Delta}) < \lambda_{\min}(\mathbf{\Sigma}^*) \end{aligned}$$

being $\rho()$ and $\lambda_{\min}()$ the spectral radius and the minimum eigenvalue of a matrix. Thus, we can consider interval covariance matrix as the set of perturbed covariance matrices $\{\mathbf{\Sigma} \pm \mathbf{\Delta}'\}$ with $0 \leqslant \mathbf{\Delta}' \leqslant \mathbf{\Delta}$ and then $\|\mathbf{\Delta}'\|_F^2 \leqslant \|\mathbf{\Delta}\|_F^2 = r^2$ being $\|A\|_F^2 = trace(AA^T)$ the Euclidean norm applied to matrices that is usually called the Frobenius norm. With these conditions, next results give upper bounds for the Kullback–Leibler divergences in Proposition 4.

Proposition 4. *Let us consider perturbations in the submatrices of the original covariance matrix with interval covariance matrices. Let* $f(x) = \ln(1+x) - \frac{x}{1+x}$, *then*

1. *If a perturbation* $\mathbf{\Delta_{YY}}$ *is added to* $\mathbf{\Sigma_{YY}}$ *satisfying next condition,*

$$\rho(\mathbf{\Delta_{YY}}) < \min\{\lambda_{\min}(\mathbf{\Sigma_{YY}}), \quad \lambda_{\min}(\mathbf{\Sigma^{Y|E}})\} \tag{8}$$

then, the perturbed covariance matrices $\mathbf{\Sigma_{YY}} + \mathbf{\Delta_{YY}}$ *and* $\mathbf{\Sigma}^{\mathbf{Y|E},\mathbf{\Delta_{YY}}}$ *are positive definite. Also it can be shown*

$$KL^{\Sigma_{YY}}(f^{\Sigma_{YY}}|f) \leqslant \frac{1}{2}\dim(\mathbf{Y})\max\{f(\lambda_{\max}(\mathbf{\Delta_{YY}}(\mathbf{\Sigma^{Y|E}})^{-1})), \quad f(\lambda_{\min}(\mathbf{\Delta_{YY}}(\mathbf{\Sigma^{Y|E}})^{-1}))\}.$$

2. *If a perturbation* $\mathbf{\Delta_{EE}}$ *is added to* $\mathbf{\Sigma_{EE}}$ *satisfying next conditions*

$$\begin{aligned} &\rho(\mathbf{\Delta_{EE}}) < \lambda_{\min}(\mathbf{\Sigma_{EE}}), \text{ and} \\ &\rho(\mathbf{\Delta^*_{EE}}) < \lambda_{\min}(\mathbf{\Sigma^{Y|E}}) \end{aligned} \tag{9}$$

with $\mathbf{\Delta_{EE}}^* = B_{Y|E}\mathbf{\Delta_{EE}}(\mathbf{\Sigma_{EE}} + \mathbf{\Delta_{EE}})^{-1}\mathbf{\Sigma_{EY}}$ *then, the perturbed covariance matrices* $\mathbf{\Sigma_{EE}} + \mathbf{\Delta_{EE}}$ *and* $\mathbf{\Sigma}^{\mathbf{Y|E},\mathbf{\Delta_{EE}}}$ *are positive definite. Also it can be shown*

$$KL^{\Sigma_{EE}}(f^{\Sigma_{EE}}|f) \leqslant \frac{1}{2}\dim(\mathbf{Y})\max\{f(\lambda_{\max}(\mathbf{\Delta^*_{EE}}(\mathbf{\Sigma^{Y|E}})^{-1})), \quad f(\lambda_{\min}(\mathbf{\Delta^*_{EE}}(\mathbf{\Sigma^{Y|E}})^{-1}))\} + \frac{1}{2}\left[(\delta^*_{\mathbf{EE}})^T(\mathbf{\Sigma^{Y|E}} + \mathbf{\Delta^*_{EE}})^{-1}(\delta^*_{\mathbf{EE}})\right]$$

for $\delta_{\mathbf{EE}}^* = B_{Y|E}\mathbf{\Delta_{EE}}(\mathbf{\Sigma_{EE}} + \mathbf{\Delta_{EE}})^{-1}(\mathbf{e} - \mu_{\mathbf{E}})$.

3. *If a perturbation* $\mathbf{\Delta_{YE}}$ *is added to* $\mathbf{\Sigma_{YE}}$ *satisfying next condition*

$$\rho(\mathbf{\Delta^*_{YE}}) < \lambda_{\min}(\mathbf{\Sigma^{Y|E}}) \tag{10}$$

with $\mathbf{\Delta^*_{YE}} = -(\mathbf{\Delta_{YE}}B^T_{Y|E} + B_{Y|E}\mathbf{\Delta}^T_{YE} + \mathbf{\Delta_{YE}}\mathbf{\Sigma}^{-1}_{\mathbf{EE}}\mathbf{\Delta_{EY}})$, *then* $\mathbf{\Sigma}^{\mathbf{Y|E},\mathbf{\Delta_{YE}}}$ *is positive definite and*

$$KL^{\Sigma_{YE}}(f^{\Sigma_{YE}}|f) \leqslant \frac{1}{2}\dim(\mathbf{Y})\max\left\{f\left(\lambda_{\max}\left(\mathbf{\Delta^*_{YE}}(\mathbf{\Sigma^{Y|E}})^{-1}\right)\right), \quad f\left(\lambda_{\min}\left(\mathbf{\Delta^*_{YE}}(\mathbf{\Sigma^{Y|E}})^{-1}\right)\right)\right\} + \frac{1}{2}\left[(\delta^*_{\mathbf{YE}})^T(\mathbf{\Sigma^{Y|E}} + \mathbf{\Delta^*_{YE}})^{-1}(\delta^*_{\mathbf{YE}})\right]$$

for $\delta_{\mathbf{YE}}^* = \mathbf{\Delta_{YE}}(\mathbf{\Sigma_{EE}})^{-1}(\mathbf{e} - \mu_{\mathbf{E}})$.

With the results presented in Proposition 4 it is possible to set an upper bound for the KL divergences introduced in Proposition 3, given certain conditions shown in this Proposition 4.

If there exist some inaccuracies in the parameters describing a GBN, it is possible to carry out the proposed sensitivity analysis using the expressions given in Propositions 1 and 3.

If we compare the expressions obtained in both propositions we see that the expressions in Proposition 1 only depend on changes in the mean vector. The first expression of Proposition 3 only depends on changes in the covariance matrix. Finally, in the last two expressions of Proposition 3, mean, variances and covariances are perturbed. Therefore both expressions are similar and depend on both the mean vector and the covariance matrix. In all cases, the considered parameters are the perturbed ones after the evidence propagation.

Then, by computing the KL divergence for each case, we can determine which set of parameters must be reviewed to describe the network more accurately.

Propositions 2 and 4 show upper bounds for the expressions computed in Propositions 1 and 3.

4.3. Uncertainty about the mean vector and the covariance matrix

If in Propositions 1 and 3 all the divergences are close to zero, it is possible to conclude that the network is not sensitive to the proposed perturbations. For this reason we have developed a simple method of analyzing robustness as well. The method is based on comparing the original network output after evidence propagation, given by the conditional probability density f of the distribution $\mathbf{Y}|\mathbf{E} \sim N(\mu^{\mathbf{Y}|\mathbf{E}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$, to the perturbed network output, given by the conditional probability density $f^{\delta,\mathbf{\Delta}}$ of the distribution $\mathbf{Y}|\mathbf{E}, \delta, \mathbf{\Delta} \sim N(\mu^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}})$. Thus, for this analysis only one perturbed model is considered. In this case the KL divergence yields

$$\begin{aligned} KL(f^{\delta,\Delta}|f) &= E_f\left[\ln\frac{f}{f^{\delta,\Delta}}\right] \\ &= \frac{1}{2}\left[-\ln|\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}})^{-1}| + trace\left(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}})^{-1}\right) - \dim(\mathbf{Y})\right] + \frac{1}{2}\left[(\mu^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}} - \mu^{\mathbf{Y}|\mathbf{E}})^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}})^{-1}(\mu^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}} - \mu^{\mathbf{Y}|\mathbf{E}})\right] \end{aligned} \tag{11}$$

By computing this expression, we can decide whether or not the GBN is robust under the considered perturbations. McCulloch [24] introduces a calibration for the KL divergence that can be used as a guide to evaluate the value obtained for the KL measure. Nevertheless, if the KL divergence is very close to zero, then the network is robust to these perturbations; otherwise, it is necessary to review the parameters assigned to the quantitative part of the network in more detail. If we consider interval covariance matrices for perturbations,

$$\begin{aligned} \mu^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}} &= \mu^{\mathbf{Y}|\mathbf{E}} + \delta^*(\delta, \mathbf{\Delta}) \\ \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E},\delta,\mathbf{\Delta}} &= \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}} + \mathbf{\Delta}^*(\mathbf{\Delta}) \end{aligned}$$

then, similar results as those obtained previously for the upper bounds can be shown.

The sensitivity and robustness analyses introduced in this section are implemented with a polynomial algorithm that computes the divergences from a set of uncertain parameters.

5. Application of results

We present two different examples, which make different assumptions about the degree of uncertainty. In both examples the network is built and an expert, who has not been part of the process of building it, disagrees with some of the parameters describing the model.

Example 3 shows some of the uncertainties more decisive for the expert that disagrees with the network. Example 4, in addition to those uncertainties of Example 3, is regards more uncertainties in the parameters. Then, all the uncertainties listed in Example 3 also appear in Example 4 together with some other more. Therefore, it is expected that the network's output of Example 4 is more sensitive than the network's output of Example 3 for uncertainty introduced because all uncertainty in Example 3 is in Example 4 and that the analyses proposed reflect this situation.

These perturbations are proposed by the expert as could have chosen any other perturbation given its experience. The only one restriction is to keep the covariance matrix positive definite before the evidence propagation; this is the only restriction on the setting of $\mathbf{\Delta}$, while for values of δ there is no restriction. Then, the examples introduced in this Section are characterized by perturbations collected in Example 3 and they are also in Example 4 together with some further perturbations. Moreover, for all proposed perturbations to the covariance matrix, $\mathbf{\Delta}$, the perturbed matrices obtained after adding the partition of $\mathbf{\Delta}$ in the covariance matrix, must be positive definite.

Example 3. For the GBN in Example 1, assume that experts disagree on the values of some parameters. For example, the mean of X_5 could be either 4 or 5. They also offer different opinions about the variances of X_2, X_3 and X_5, and about the covariance between X_1 and X_3. When we quantify these uncertainties, the perturbed mean vector δ and the perturbed covariance matrix $\mathbf{\Delta}$ are

$$\delta = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}; \quad \mathbf{\Delta} = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Again, the set of variables of interest is $\mathbf{Y} = (X_3, X_4, X_5)$ and the set of evidential variables is $\mathbf{E} = \{X_1 = 2, X_2 = 4\}$.

As can be seen, there is uncertainty about some of the parameters describing the network. This uncertainty is reflected in the vector δ and the matrix $\mathbf{\Delta}$. Then, we can now study the sensitivity of the network given the uncertainty δ and $\mathbf{\Delta}$. We set

$$\delta_{\mathbf{Y}} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}; \quad \delta_{\mathbf{E}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{\Delta}_{\mathbf{YY}} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}; \quad \boldsymbol{\Delta}_{\mathbf{EE}} = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}; \quad \boldsymbol{\Delta}_{\mathbf{YE}} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The methodology proposed let us to evaluate different sets of uncertainty using the expressions shown in Propositions 1 and 3. With the obtained results, it is possible to compare different uncertainty situations and determine the set of parameters that must be described with great precision.

Then, next results are obtained

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) = 0.167$$
$$KL^{\mu_{\mathbf{E}}}(f^{\mu_{\mathbf{E}}}|f) = 0$$
$$KL^{\Sigma_{\mathbf{YY}}}(f^{\Sigma_{\mathbf{YY}}}|f) = 0.338$$
$$KL^{\Sigma_{\mathbf{EE}}}(f^{\Sigma_{\mathbf{EE}}}|f) = 0.203$$
$$KL^{\Sigma_{\mathbf{YE}}}(f^{\Sigma_{\mathbf{YE}}}|f) = 0.447$$

Note that all the divergences are not too large, in the sense of [24]. Then with the uncertainty proposed we conclude the network is not very sensitive to the proposed block perturbations.

Notice that $KL^{\mu_{\mathbf{E}}}(f^{\mu_{\mathbf{E}}}|f) = 0$, because there are no perturbations added to $\mu_{\mathbf{E}}$. Thus, $\delta_{\mathbf{E}}$ is zero.

With this sensitivity analysis it seems that the network is rather robust under the proposed perturbations. This can be checked by computing expression (11). Then, the KL divergence is computed for the whole set of perturbations and therefore the robustness of the network to these global perturbation can be studied.

$$KL(f^{\delta,\Delta}|f) = 0.857$$

The divergence that determines the robustness of the model is small and consequently the network is considered robust for uncertainty of the example.

We can also compare the effect of different individual perturbations

$$\delta_{\mathbf{Y}} = \begin{pmatrix} \delta_{X_3} \\ \delta_{X_4} \\ \delta_{X_5} \end{pmatrix}; \quad \delta_{\mathbf{E}} = \begin{pmatrix} \delta_{X_1} \\ \delta_{X_2} \end{pmatrix}$$

making calculations for each δ_{X_i} from -1 to 1 by 0.001 when the remaining variables are not perturbed. Fig. 2 illustrates the resulting divergences with large values for perturbations in X_1 and rather small for X_5 perturbations.

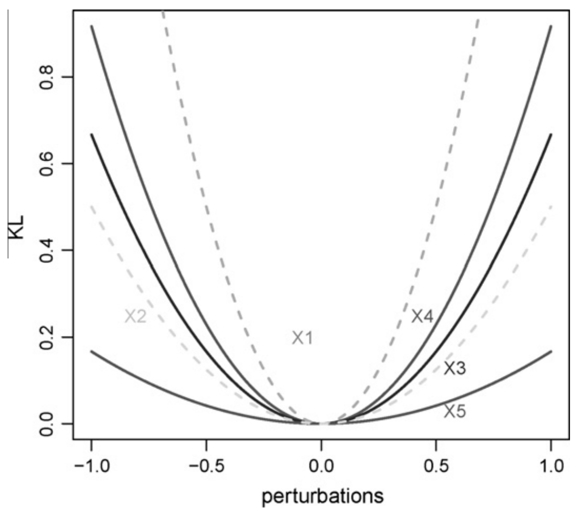

Fig. 2. KL^{μ} divergence measures to perturbations in the individual components of the mean vector of Example 1.

Using the upper bounds results for perturbations such that $\|\delta'_{\mathbf{Y}}\|_F^2 \leqslant 1$ and the minimum eigenvalue of $\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}$ given by 0.393314, we have

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) \leqslant \frac{1}{2(0.3933140)} = 1.2712$$

where the maximum is attained with $\delta^*_{\mathbf{Y}} = \begin{pmatrix} -0.52 \\ -0.79 \\ 0.32 \end{pmatrix}$, the corresponding eigenvector.

Perturbing the submatrices of the covariance matrix with symmetric interval matrix we obtain positive definite covariance matrices after perturbations. This condition restrict perturbations to matrices with small values but preserving the structure of the example perturbations.

In particular, the condition (8) holds for

$$\mathbf{\Delta}^P_{\mathbf{YY}} = 0.18\mathbf{\Delta}_{\mathbf{YY}} = \begin{pmatrix} 0.36 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.36 \end{pmatrix}$$

then for $0 \leqslant \mathbf{\Delta}'_{YY} \leqslant \mathbf{\Delta}^P_{YY}$ with $\|\mathbf{\Delta}'_{YY}\|_F^2 \leqslant \|\mathbf{\Delta}^P_{YY}\|_F^2 = (0.51)^2$ we have

$$KL^{\mathbf{\Sigma}_{\mathbf{YY}}}(f^{\mathbf{\Sigma}_{\mathbf{YY}}}|f) \leqslant 0.11.$$

As for the condition (9) the particular perturbation

$$\mathbf{\Delta}^P_{\mathbf{EE}} = 0.17\begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0.51 \end{pmatrix}$$

then for $0 \leqslant \mathbf{\Delta}'_{EE} \leqslant \mathbf{\Delta}^P_{EE}$ with $\|\mathbf{\Delta}'_{EE}\|_F^2 \leqslant \|\mathbf{\Delta}^P_{EE}\|_F^2 = (0.51)^2$ we have

$$KL^{\mathbf{\Sigma}_{EE}}(f^{\mathbf{\Sigma}_{\mathbf{EE}}}|f) \leqslant 0.093.$$

Finally, given that the matrix $\mathbf{\Delta}^*_{\mathbf{YE}} = \mathbf{\Delta}_{\mathbf{YE}}B^T_{Y|E} + B_{Y|E}\mathbf{\Delta}^T_{YE} + \mathbf{\Delta}_{\mathbf{YE}}\mathbf{\Sigma}^{-1}_{\mathbf{EE}}\mathbf{\Delta}_{\mathbf{EY}}$ has to satisfy condition (10) we can use

$$\mathbf{\Delta}_{\mathbf{YE}} = \begin{pmatrix} -0.08 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and the corresponding

$$\mathbf{\Delta}^*_{\mathbf{YE}} = \begin{pmatrix} 0.318 & 0 & 0.16 \\ 0 & 0 & 0 \\ 0.16 & 0 & 0 \end{pmatrix}$$

with $\|\mathbf{\Delta}^*_{\mathbf{YE}}\|_F^2 = (0.392)^2$, to obtain

$$KL^{\mathbf{\Sigma}_{\mathbf{YE}}}(f^{\mathbf{\Sigma}_{YE}}|f) \leqslant 0.059$$

With respect to the calibration of the KL divergences, the Table 1 in [24] could be considered.

Example 4. Working with the GBN given in Example 1, the experts now disagree on the values of more parameters. We introduce uncertainty into the parameters that describe X_4 and X_2, so that the perturbed mean vector δ and perturbed covariance matrix $\mathbf{\Delta}$ are

$$\delta = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \\ -1 \end{pmatrix}; \quad \mathbf{\Delta} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & 0 & 1 \\ -1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 & 2 \end{pmatrix}$$

The variables of interest are still $\mathbf{Y} = (X_3, X_4, X_5)$, and the evidential variables $\mathbf{E} = \{X_1 = 2, X_2 = 4\}$.

In this example the uncertainty of the parameters describing the network is larger than in the previous example. Now there is uncertainty in all parameter sets (remember that in the above example there is no uncertainty associated with the mean of the evidential variables). Again, we use the methodology proposed in this paper. Thus, to perform the proposed sensitivity analysis, the new partitions are

$$\delta_{\mathbf{Y}} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}; \quad \delta_{\mathbf{E}} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\mathbf{\Delta_{YY}} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 2 \end{pmatrix}; \quad \mathbf{\Delta_{EE}} = \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}; \quad \mathbf{\Delta_{YE}} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Computing expressions in Propositions 1 and 3, the KL divergences obtained are

$$KL^{\mu_\mathbf{Y}}(f^{\mu_\mathbf{Y}}|f) = 1.75$$

$$KL^{\mu_\mathbf{E}}(f^{\mu_\mathbf{E}}|f) = 2$$

$$KL^{\Sigma_\mathbf{YY}}(f^{\Sigma_\mathbf{YY}}|f) = 0.49$$

$$KL^{\Sigma_\mathbf{EE}}(f^{\Sigma_\mathbf{EE}}|f) = 0.383$$

$$KL^{\Sigma_\mathbf{YE}}(f^{\Sigma_\mathbf{YE}}|f) = 1.888$$

To evaluate the obtained values for the KL measure, McCulloch [24] introduces a calibration for the KL divergence. As can be seen, in this example divergences are larger than in the previous one. If the KL divergence is close to zero the network is not sensitive to these perturbations, otherwise experts should review the literature and databases again to be more precise when defining the parameters.

In this case, some divergences are larger than 1. We can therefore conclude that the network is very sensitive to the proposed perturbations. Then, the parameters of the mean vector and the covariances between interest and evidential variables should be reviewed to describe the GBN more accurately.

When evaluating the robustness of the network, we obtain a very large value for the total KL divergence:

$$KL(f^{\delta,\Delta}|f) = 17.022$$

For a better understanding of the problem, can be made some more general perturbations. Let

$$\delta_\mathbf{Y} = \begin{pmatrix} 0 \\ \delta_{X_4} \\ \delta_{X_5} \end{pmatrix}; \quad \delta_\mathbf{E} = \begin{pmatrix} \delta_{X_1} \\ \delta_{X_2} \end{pmatrix}$$

with $\|\delta_\mathbf{Y}\|_F^2 \leqslant 2$ and $\|\delta_\mathbf{E}\|_F^2 \leqslant 4$. The calculations of the KL measure for simultaneous perturbations in an extremely fine grids for $(\delta_{X_4}, \delta_{X_5})$ and $(\delta_{X_1}, \delta_{X_2})$ are presented in Figs. 3 and 4 respectively. Also it is obtained the maximum divergence for $\delta_\mathbf{Y}$ perturbations in the grid is 2.0868, attained for $\delta_{X_4} = 1.3228$ and $\delta_{X_5} = -0.5$. For $\delta_\mathbf{E}$ perturbations in the grid, the largest deviation is 9.9994 for $\delta_{X_1} = 1.792, \delta_{X_2} = 0.904$.

To explore $\mathbf{\Delta_{YY}}$ perturbations satisfying the positive definiteness conditions required, we will use a uniform discrete grid of $\delta_{var(X_3)}$ and $\delta_{var(X_5)}$ values from 0 to 2 by 0.005 representing perturbations in the variances of X_3 and X_5 with their KL measures in Fig. 5. Also a grid for perturbations in the variance of X_4 and the covariance of X_4 and X_5 and the KL values are shown in Fig. 6.

Also, $\mathbf{\Delta_{EE}}$ perturbations can be studied in some general scenarios. Then, a grid for perturbations in the variances of X_1 and X_2 from 0 to 1 by 0.001 with the KL measures are illustrated in Fig. 7 and for perturbations in the variance of X_2 from 0 to 3 by 0.005 and in the covariance of X_1 and X_2 from 0 to 1.5 by 0.001 in Fig. 8. In this last graphic, a very large increase is observed for perturbations in the region $\delta_{cov(X_1,X_2)} > 1$ and $\delta_{var(X_2)}$ close to zero.

Finally, the study of $\mathbf{\Delta_{YE}}$ perturbations only for some small values that result in positive definite covariance matrices is presented in Fig. 9 with perturbations in the covariance of X_3 and X_1 and in the covariance of X_5 and X_2 simultaneously. Some interesting effects are observed with respect to the robustness of the network from this type of uncertainty.

Summing up, the KL divergence that evaluates the robustness of the network is not small. Again, we can conclude that some parameters must be reviewed to describe the network more accurately, as we determined with the sensitivity analysis. Specifically, mean vector uncertainty introduce a large fluctuation in the network output that is enlarged if we add some other parameter perturbations.

If we study upper bounds in this case we have:

1. For general δ'_Y perturbations such that $\|\delta'_\mathbf{Y}\|_F^2 \leqslant \|\delta_\mathbf{Y}\|_F^2 = 2$,

   $KL^{\mu_\mathbf{Y}}(f^{\mu_\mathbf{Y}}|f) \leqslant 2.542$

   and the maximum is attained in $\begin{pmatrix} -0.741 \\ -1.118 \\ 0.449 \end{pmatrix}$
2. In this example, for δ'_E perturbations such that $\|\delta'_\mathbf{E}\|_F^2 \leqslant \|\delta_\mathbf{E}\|_F^2 = 4$ and $\|B_{Y|E}\delta'_\mathbf{E}\|_2^2 \leqslant \|B_{Y|E}\delta_\mathbf{E}\|_2^2 = 8$ the upper bound is

   $KL^{\mu_E}(f^{\mu_\mathbf{E}}|f) \leqslant 10.17.$
3. For $0 \leqslant \mathbf{\Delta}'_{YY} \leqslant \mathbf{\Delta}^P_{YY}$ with $\|\mathbf{\Delta}'_{YY}\|_F^2 \leqslant \|\mathbf{\Delta}^P_{YY}\|_F^2 = (0.41)^2$ we have

$KL^{\Sigma_{\mathbf{YY}}}(f^{\Sigma_{\mathbf{YY}}}|f) \leqslant 0.035$

with $\mathbf{\Delta}^P_{\mathbf{YY}} = 0.1\mathbf{\Delta}_{\mathbf{YY}} = \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.1 & 0.2 \\ 0 & 0.2 & 0.2 \end{pmatrix}$.

4. For $0 \leqslant \mathbf{\Delta}'_{EE} \leqslant \mathbf{\Delta}^P_{EE}$ with $\|\mathbf{\Delta}'_{EE}\|^2_F \leqslant \|\mathbf{\Delta}^P_{EE}\|^2_F = (0.3)^2$ we have

   $KL^{\Sigma_{EE}}(f^{\Sigma_{\mathbf{EE}}}|f) \leqslant 0.134$

   with $\mathbf{\Delta}^P_{\mathbf{EE}} = 0.09\mathbf{\Delta}_{\mathbf{EE}} = \begin{pmatrix} 0 & 0.09 \\ 0.09 & 0.27 \end{pmatrix}$
5. And for the last case, if we have $\|\mathbf{\Delta}^{**}_{\mathbf{YE}}\|^2_F = (0.445)^2$, it holds

   $KL^{\Sigma_{\mathbf{YE}}}(f^{\Sigma_{YE}}|f) \leqslant 0.088$

   with $\mathbf{\Delta}^P_{\mathbf{YE}} = 0.095\mathbf{\Delta}_{\mathbf{YE}} = \begin{pmatrix} -0.095 & 0 \\ 0 & 0 \\ 0 & 0.095 \end{pmatrix}$. Again the mean vector perturbations attain the larger bounds.

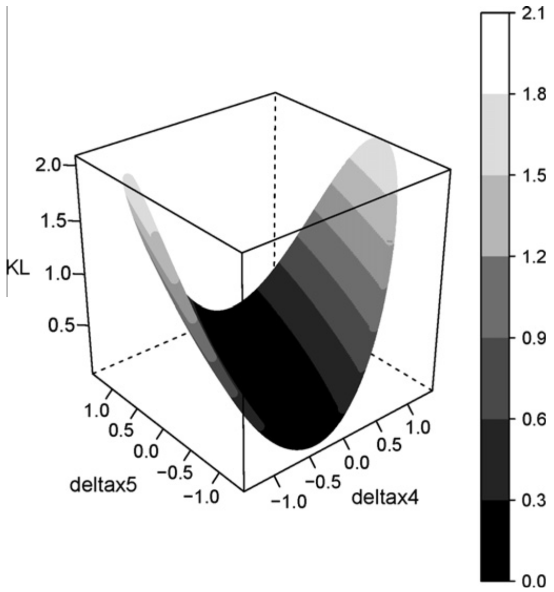

Fig. 3. $KL^{\mu_{\mathbf{Y}}}$ divergence measures to simultaneous perturbations in the means of the variables of interest X_4 and X_5 of Example 1 such that $|\delta_{\mathbf{Y}}\|^2_2 \leqslant 2$.

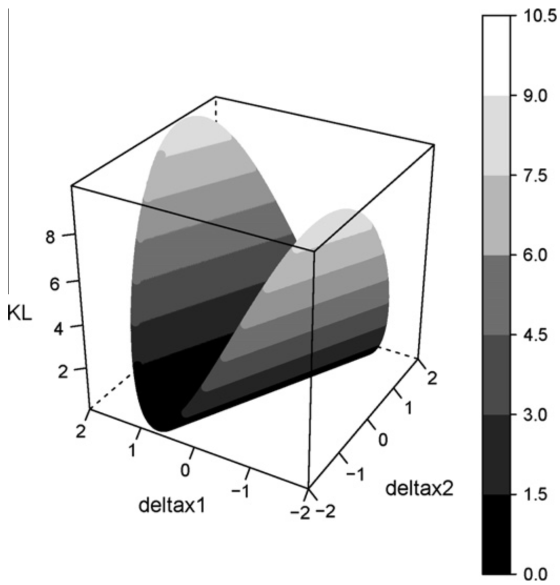

Fig. 4. $KL^{\mu_{\mathbf{E}}}$ divergence measures to simultaneous perturbations in the means of the evidential variables X_1 and X_2 of Example 1 such that $\|\delta_{\mathbf{E}}\|^2_2 \leqslant 4$.

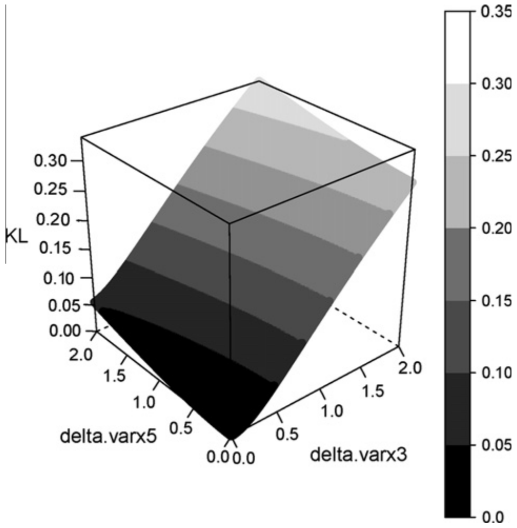

Fig. 5. $KL^{\Sigma_{YY}}$ divergence measures to simultaneous perturbations in the variances of X_3 and X_5 of the block $\mathbf{\Sigma_{YY}}$ of Example 1.

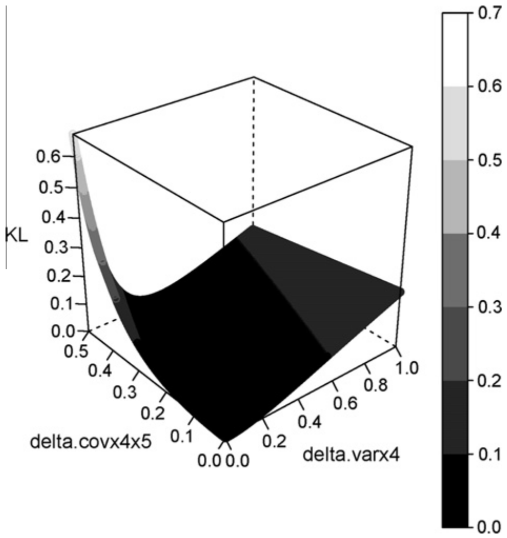

Fig. 6. $KL^{\Sigma_{YY}}$ divergence measures to simultaneous perturbations in the covariance of X_4 and X_5 and the variance of X_4 of the block $\mathbf{\Sigma_{YY}}$ of Example 1.

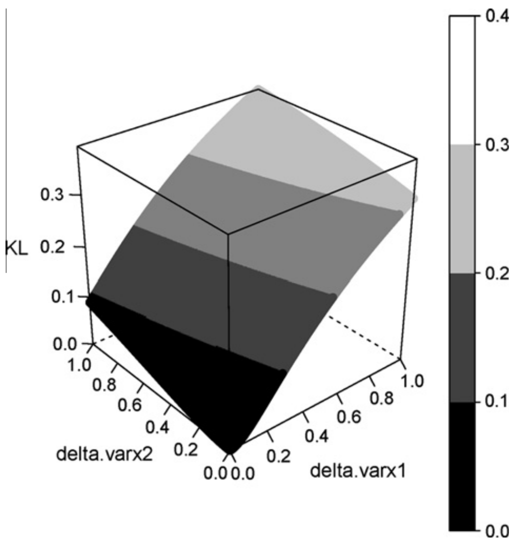

Fig. 7. $KL^{\Sigma_{EE}}$ divergence measures to simultaneous perturbations in the variances of X_1 and X_2 of the block $\mathbf{\Sigma_{EE}}$ of Example 1.

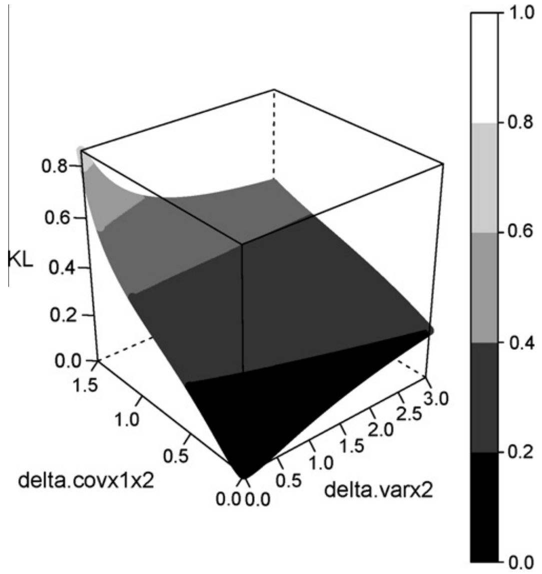

Fig. 8. $KL^{\Sigma_{EE}}$ divergence measures to simultaneous perturbations in the variance of X_2 and the covariance of X_1 and X_2 of the block $\mathbf{\Sigma_{EE}}$ of Example 1.

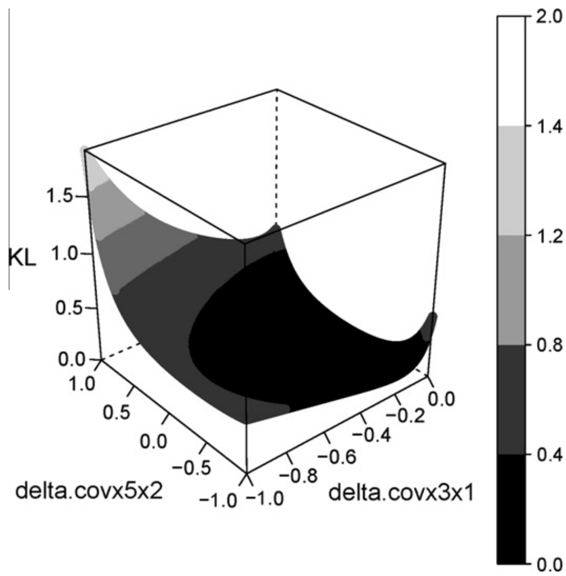

Fig. 9. $KL^{\Sigma_{YE}}$ divergence measures to simultaneous perturbations the covariance of X_3 and X_1 and the covariance of X_5 and X_2 of the block $\mathbf{\Sigma_{YE}}$ of Example 1.

As seen in previous examples, the methodology proposed in this paper is useful to determine the parameters that must be described accurately in a Gaussian Bayesian network. Furthermore, it is easy to use and allows to describe the network accurately.

6. Case study: damage assessment of a reinforced concrete structure

In this section, we study a complex problem (see [3]). Its objective is to assess the extent of damage to a building's reinforced concrete structures. Other authors [23] use this example to study sensitivity to kurtosis deviations from Gaussianity.

As shown in Table 1, 24 variables are necessary to describe the problem. Variables $X_{17},\ldots,X_{23}$ are intermediate, unobservable variables defining some partial states of the structure. It is assumed that an expert can examine a given concrete beam to sequentially obtain the values of $X_1,\ldots,X_{16}$. We consider $\mathbf{E} = (X_1,\ldots,X_{16})$ to be evidential variables, and $\mathbf{Y} = (X_{17},\ldots,X_{24})$ to be variables of interest.

The DAG that represents the dependence structure of the variables is shown in Fig. 10. The mean of every variable is zero. The covariance matrix is computed using the coefficients β_{ji} shown in the edges of Fig. 10. The conditional variances are set to 1 for observable variables, and to 10^{-4} for other variables.

The evidence proposed is $\mathbf{E} = \{X_1 = 1,\ldots,X_{16} = 1\}$. A complete description of this model and its evidence is given in [3].

There is uncertainty in some means and conditional variances of the network. Two of the evidential variables, X_{15} and X_{16}, are close to X_{24} in value. The variable of interest X_{24} is also defined with uncertainty. It is thus plausible to define a perturbed

Table 1
Description of the case study variables.

X_i	Description
X_1	Corrosion
X_2	Shrinkage
X_3	Number of flexure cracks
X_4	Number of shear cracks
X_5	Content of chlorine in the air
X_6	pH value in the air
X_7	Humidity
X_8	Structure age
X_9	Cover
X_{10}	Length of the worst flexure cracks
X_{11}	Breadth of the worst flexure crack
X_{12}	Position of the worst flexure crack
X_{13}	Breadth of the worst shear crack
X_{14}	Position of the worst shear crack
X_{15}	Deflection of the beam
X_{16}	Weakness of the beam
X_{17}	Corrosion state
X_{18}	Worst cracking in flexure domain
X_{19}	Shrinkage cracking
X_{20}	Cracking state in flexure domain
X_{21}	Steel corrosion
X_{22}	Cracking state in shear domain
X_{23}	Cracking state
X_{24}	Damage assessment

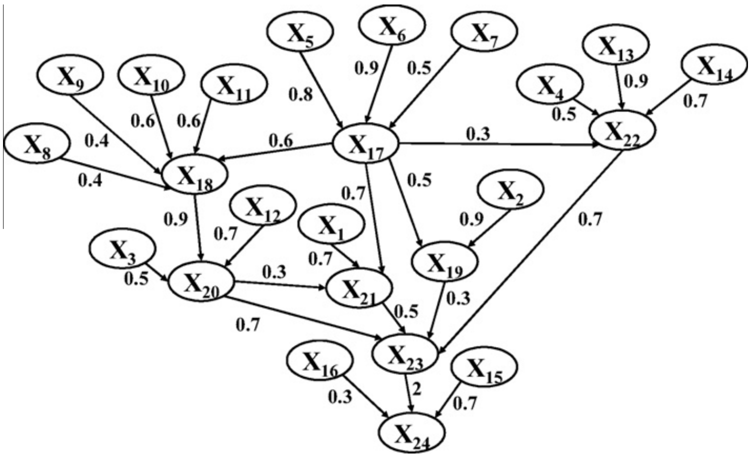

Fig. 10. DAG for the damage assessment of reinforced concrete structure with β_{ji} values.

model where the means of X_{15}, X_{16} and X_{24} are set to 1 instead of zero. The conditional variances of these variables are set to 1.5 instead of 1 (remember that X_{24} is an observable variable of interest) then we have a perturbation $\mathbf{\Delta}_D$ for D.

To study the sensitivity and robustness of this network, first it is necessary to compute the joint probability distribution of the perturbed model. Using (3) the perturbation of the covariance matrix, $\mathbf{\Delta}$, is obtained from

$$\mathbf{\Sigma} + \Delta = [(\mathbf{I} - \mathbf{B})^{-1}]^T(\mathbf{D} + \mathbf{\Delta_D})(\mathbf{I} - \mathbf{B})^{-1}$$

and then it is partitioned as follows:

$$\mathbf{\Delta} = \begin{pmatrix} \mathbf{\Delta_{YY}} & \mathbf{\Delta_{YE}} \\ \mathbf{\Delta_{EY}} & \mathbf{\Delta_{EE}} \end{pmatrix}$$

and with the previous results we have,

$$KL^{\mu_\mathbf{Y}}(f^{\mu_\mathbf{Y}}|f) = 0.5$$
$$KL^{\mu_\mathbf{E}}(f^{\mu_\mathbf{E}}|f) = 0.5$$
$$KL^{\mathbf{\Sigma_{YY}}}(f^{\mathbf{\Sigma_{YY}}}|f) = 0.0704$$
$$KL^{\mathbf{\Sigma_{EE}}}(f^{\mathbf{\Sigma_{EE}}}|f) = 0.0539$$
$$KL^{\mathbf{\Sigma_{YE}}}(f^{\mathbf{\Sigma_{YE}}}|f) = 1.1272$$

On the basis of these values, we conclude that the proposed uncertainties do not disturb too much the network output after evidence propagation. Nevertheless, the parameters that describe covariances between evidential and interest variables should be examined more carefully, because the divergence measure $KL^{\Sigma_{YE}}(f^{\Sigma_{YE}}|f)$ is over 1.

To complete this example, we study the robustness of the network under the proposed uncertainties. The KL divergence shows that the network is robust to these perturbations because the KL divergence has decreased to

$$KL(f^{\delta,\Delta}|f) = 0.03606$$

In fact, the global effect of different parameter perturbations is smaller than the particular deviations that can be explained if viewed as the result of opposing effects. Some perturbations $\delta_{\mathbf{X}_{15}}$ and $\delta_{\mathbf{X}_{16}}$, for the means of X_{15} and X_{16}, in a grid with values from -1 to 1 by 0.005 are considered and the corresponding KL measures presented in Fig. 11. Also we can study the effect of d perturbation same in all the conditional variances of X_{15}, X_{16} and X_{24} with d ranging from 0 to 0.5 by 0.05 and make comparisons on the divergences related to perturbations in the different blocks as well as in the total covariance matrix as it is shown in Fig. 12. As expected the highest values correspond to induced $\mathbf{\Sigma_{YE}}$ perturbations and the global KL is always below.

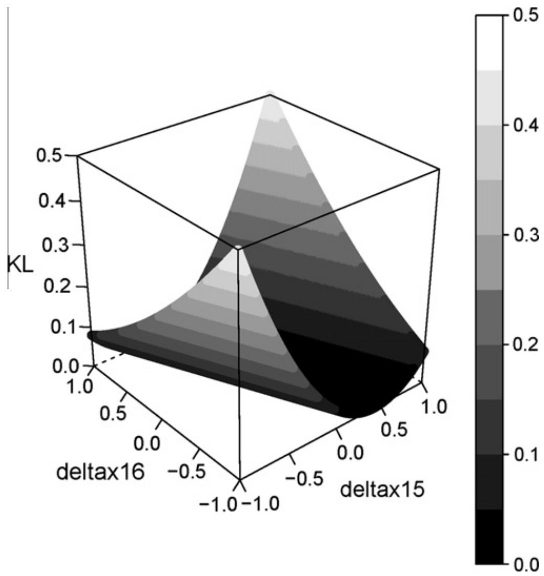

Fig. 11. KL^{μ_E} divergence measures to simultaneous perturbations in the means of the evidential variables X_{15} and X_{16} of Case Study such that $\|\delta_E\|_2^2 \leqslant 2$.

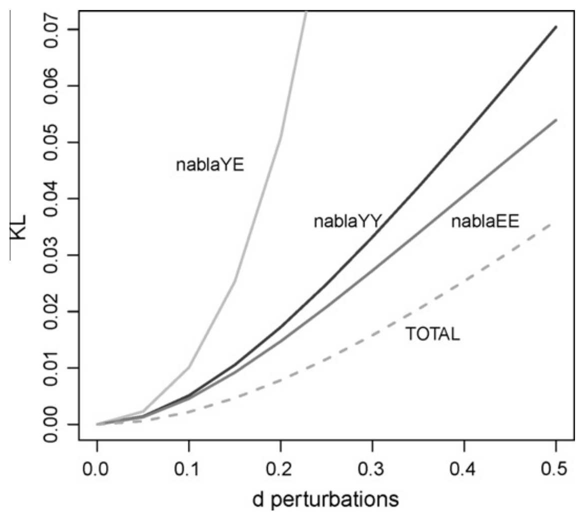

Fig. 12. KL divergence measures to perturbations in the different blocks by d perturbation to be the same in all the conditional variances of X_{15}, X_{16} and X_{24}.

Thus, we can say our proposed method is easy to calculate and apply even for complicated problems. It also allows further study of the effects that uncertain parameters can have on the network output. Since then some of the decisions necessary to define more accurately the most influential parameters can be taken.

7. Conclusions

This paper has presented a new method for analyzing the sensitivity and robustness of a GBN when some of the parameters describing the quantitative part of the network are inaccurate or uncertain. The KL divergence is used to compare the overall behaviors of the original network and a perturbed network, thereby evaluating the effect of changes in the network parameters after evidence propagation for the variables of interest. The proposed method is simple to calculate for any GBN, and can handle any kind of perturbation or inaccuracy in the network parameters. Then, it is possible to study either large or small perturbations in the uncertain parameters of the network.

The sensitivity analysis considers different sets of parameters, depending on which kinds of variables are perturbed (interest or evidential), and whether the uncertainties lie in their mean vector or covariance matrix. Each case yields a different divergence, making it possible to know which set of variables most strongly disturbs the network output after evidence propagation. When the divergences obtained by the sensitivity analysis are small, the network may be considered robust under the proposed perturbations. To confirm this result, we also develop a robustness analysis method that evaluates all the perturbations at the same time. There still is not an answer when KL divergence is sufficiently large. We consider that until that divergence is not used more in practical contexts with different uncertainty situations an accurate answer to that question cannot be given.

Acknowledgements

This research has been supported by the Spanish Ministerio de Ciencia e Innovación, Grant MTM 2008-03282 and partially by GR58/08-A, 910395 – Métodos Bayesianos by the BSCH-UCM from Spain.

We also want to thank the referees for their useful comments.

Appendix A

Proof of Proposition 1. After the evidence propagation, the output of the original network is $\mathbf{Y}|\mathbf{E} \sim N(\mu^{\mathbf{Y}|\mathbf{E}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$ and parameters of the perturbed models can be obtained directly taking into account the evidence propagation process.

1. The output of the perturbed model is $\mathbf{Y}|\mathbf{E}, \delta_{\mathbf{Y}} \sim N(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{Y}}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$. The covariance matrix is the same in both models, original and perturbed because, as follows from (4), the conditional covariance matrix $\Sigma^{\mathbf{Y}|\mathbf{E}}$ does not depend on the mean vector, unlike the conditional mean vector with respect to the covariance matrix, then is not affected by perturbations in the mean vector and remains the same for both models regardless of the block mean vector to be perturbed. To compute $KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f)$ we have that, *trace* $(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}) = \dim(\mathbf{Y})$, then

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) = \frac{1}{2}\left[\left(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{Y}}} - \mu^{\mathbf{Y}|\mathbf{E}}\right)^T \left(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}\right)^{-1}\left(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{Y}}} - \mu^{\mathbf{Y}|\mathbf{E}}\right)\right] = \frac{1}{2}\left[\delta_{\mathbf{Y}}^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}\delta_{\mathbf{Y}}\right]$$

2. In this case, the output of the perturbed model is $\mathbf{Y}|\mathbf{E}, \delta_{\mathbf{E}} \sim N(\mu^{\mathbf{Y}|\mathbf{E},\ \delta_{\mathbf{E}}}, \mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})$. The covariance matrix is the same in both models, again. Then

$$KL^{\mu_{\mathbf{E}}}(f^{\mu_{\mathbf{E}}}|f) = \frac{1}{2}\left[\left(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{E}}} - \mu^{\mathbf{Y}|\mathbf{E}}\right)^T (\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}\left(\mu^{\mathbf{Y}|\mathbf{E},\delta_{\mathbf{E}}} - \mu^{\mathbf{Y}|\mathbf{E}}\right)\right] = \frac{1}{2}\left[\delta_{\mathbf{E}}^T\left(\mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1}\right)^T\left(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}\right)^{-1}\left(\mathbf{\Sigma}_{\mathbf{YE}}\mathbf{\Sigma}_{\mathbf{EE}}^{-1}\right)\delta_{\mathbf{E}}\right] \quad \square$$

Proof of Proposition 2. Given that

$$KL^{\mu_{\mathbf{Y}}}(f^{\mu_{\mathbf{Y}}}|f) = \frac{1}{2}[\delta_{\mathbf{Y}}^T(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}\delta_{\mathbf{Y}}]$$

and $\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}}$ is positive definite with ordered eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_q \geqslant 0$ and associated normalized eigenvectors $e_1, \ldots, e_q$, we have $(\mathbf{\Sigma}^{\mathbf{Y}|\mathbf{E}})^{-1}$ has eigenvalues $0 \leqslant \frac{1}{\lambda_1} \leqslant \cdots \leqslant \frac{1}{\lambda_q}$ and the same eigenvectors then using the Maximization Lemma for positive definite matrices [18]

$$\max_{\delta_{\mathbf{Y}}\neq 0}\frac{\delta_{\mathbf{Y}}^T(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}\delta_{\mathbf{Y}}}{\delta_{\mathbf{Y}}^T\delta_{\mathbf{Y}}}=\frac{1}{\lambda_q}$$

attained when $\delta_{\mathbf{Y}} = e_q$ and the result (6) holds.

Now we have

$$KL^{\mu_{\mathbf{E}}}(f^{\mu_{\mathbf{E}}}|f)=\frac{1}{2}\left[\delta_{\mathbf{E}}^T(\Sigma_{\mathbf{YE}}\Sigma_{\mathbf{EE}}^{-1})^T(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}(\Sigma_{\mathbf{YE}}\Sigma_{\mathbf{EE}}^{-1})\delta_{\mathbf{E}}\right]$$

and the previous Maximization Lemma for positive definite matrices is applied to $\delta_{\mathbf{Y}}^{*T}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}\delta_{\mathbf{Y}}^*$ being $\delta_{\mathbf{Y}}^* = (\Sigma_{\mathbf{YE}}\Sigma_{\mathbf{EE}}^{-1})\,\delta_{\mathbf{E}}$ to obtain (7). □

Proof of Proposition 3. After the evidence propagation, the output of the original network is $\mathbf{Y}|\mathbf{E} \sim N(\mu^{\mathbf{Y}|\mathbf{E}},\Sigma^{\mathbf{Y}|\mathbf{E}})$.

In next cases, different perturbed models are considered. Parameters of the perturbed models can be obtained directly taking into account the evidence propagation process. The KL divergence is calculated directly using the.

1. The output of the perturbed model is $\mathbf{Y}|\mathbf{E}, \Delta_{\mathbf{YY}} \sim N(\mu^{\mathbf{Y}|\mathbf{E}},\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}})$, with $\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}} = \Sigma^{\mathbf{Y}|\mathbf{E}} + \Delta_{\mathbf{YY}}$. In this case, the mean vector is the same in both models, original and perturbed. Then,

$$KL^{\Sigma_{\mathbf{YY}}}(f^{\Sigma_{\mathbf{YY}}}|f)=\frac{1}{2}\left[\ln|\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}})^{-1}|+trace\left(\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}})^{-1}\right)-\dim(\mathbf{Y})\right]$$

2. In this case, the expression associated with the KL divergence is the.
3. The KL divergence is

$$\begin{aligned}KL^{\Sigma_{\mathbf{YE}}}(f^{\Sigma_{\mathbf{YE}}}|f)&=\frac{1}{2}\left[\ln|\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}})^{-1}|+trace(\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}})^{-1})-\dim(\mathbf{Y})\right]\\&\quad+\frac{1}{2}\left[(\mu^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}}-\mu^{\mathbf{Y}|\mathbf{E}})^T(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}})^{-1}(\mu^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}}-\mu^{\mathbf{Y}|\mathbf{E}})\right]\\&=\frac{1}{2}\left[\ln|\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YY}}})^{-1}|+trace\left(\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}})^{-1}\right)-\dim(\mathbf{Y})\right]\\&\quad+\frac{1}{2}\left[(\mathbf{e}-\mu_{\mathbf{E}})^T(\Sigma_{\mathbf{EE}}^{-1})^T\Delta_{\mathbf{YE}}^T(\Sigma^{\mathbf{Y}|\mathbf{E},\Delta_{\mathbf{YE}}})^{-1}\Delta_{\mathbf{YE}}\Sigma_{\mathbf{EE}}^{-1}(\mathbf{e}-\mu_{\mathbf{E}})\right]\quad\square\end{aligned}$$

Proof of Proposition 4. Since interval covariance matrix is considered for perturbations with the conditions assumed then positive definite covariance matrices are obtained as desired. For all the cases studied it can be shown

$$\begin{aligned}\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}&=\mu^{\mathbf{Y}|\mathbf{E}}+\delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}\\\Sigma^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}&=\Sigma^{\mathbf{Y}|\mathbf{E}}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}.\end{aligned}$$

Moreover, since

$$\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}=\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E}}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}=\left(I_{\dim(Y)}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}\right)^{-1}$$

with $I_{\dim(Y)}$ the identity matrix, it follows that the first summand in (5)

$$\begin{aligned}&-\ln|\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}|+trace\left(\Sigma^{\mathbf{Y}|\mathbf{E}}(\Sigma^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}\right)-\dim(\mathbf{Y})\\&=\ln|\left(I_{\dim(Y)}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}\right)|+trace\left(\left(I_{\dim(Y)}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}\right)^{-1}-I_{\dim(Y)}\right)\\&=\sum_{i=1}^{\dim(Y)}\left(\ln(1+\lambda_i)+\frac{1}{1+\lambda_i}-1\right)=\sum_{i=1}^{\dim(Y)}\left(\ln(1+\lambda_i)-\frac{\lambda_i}{1+\lambda_i}\right).\end{aligned}$$

with $\{\lambda_i\}$ the eigenvalues of $\Delta^{\mathbf{Y}|\mathbf{E},p_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1}$. Then, the upper bound is attained with $\lambda_{\max}(\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1})$ or $\lambda_{\min}(\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}(\Sigma^{\mathbf{Y}|\mathbf{E}})^{-1})$, substituting $\Delta^{\mathbf{Y}|\mathbf{E},p_j}$ for the different types of perturbations.

The formula (5) when perturbations are added to $\Sigma_{\mathbf{EE}}$ or $\Sigma_{\mathbf{YE}}$ includes also the term with $\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}$ that can be bounded using the Maximization Lemma for positive definite matrices that gives

$$(\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}-\mu^{\mathbf{Y}|\mathbf{E}})^T(\Sigma^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}(\mu^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}-\mu^{\mathbf{Y}|\mathbf{E}})=(\delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^T(\Sigma^{\mathbf{Y}|\mathbf{E}}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})^{-1}(\delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})\leqslant\frac{|\delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j}\|_2^2}{\lambda_{\min}(\Sigma^{\mathbf{Y}|\mathbf{E}}+\Delta^{\mathbf{Y}|\mathbf{E},\mathbf{p}_j})}\quad\square$$

References

[1] M. Bednarski, W. Cholewa, W. Frid, Identification of sensitivities in Bayesian networks, Engineering Applications of Artificial Intelligence 17 (2004) 327–335.
[2] K. Burns, Bayesian inference in disputed authorship: a case study of cognitive errors and a new system for decision support, Information Sciences 176 (2006) 1570–1589.
[3] E. Castillo, U. Kjærulff, Sensitivity analysis in Gaussian Bayesian networks using a symbolic-numerical technique, Reliability Engineering and System Safety 79 (2003) 139–148.
[4] H. Chan, A. Darwiche, Sensitivity analysis in Bayesian networks: form single to multiple parameters, in: En Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, AUAI Press, Arlington, Virginia, USA, 2004, pp. 67–75.
[5] H. Chan, A. Darwiche, A distance measure for bounding probabilistic belief change, International Journal of Approximate Reasoning 38 (2) (2005) 149–174.
[6] V.M.H. Coupé, L.C. van der Gaag, J.D.F. Habbema, Sensitivity analysis: an aid for belief-network quantification, The Knowledge Engineering Review 15 (3) (2000) 215–232.
[7] V.M.H. Coupé, L.C. van der Gaag, Properties of sensitivity analysis of Bayesian belief networks, Annals of Mathematics and Artificial Intelligence 36 (2002) 323–356.
[8] R.G. Cowell, A.P. Dawid, S.L. Lauritzen, D.J. Spiegelhalter, Probabilistic Networks and Expert Systems, Springer, Barcelona, 1999.
[9] M.A. Gómez-Villegas, P. Main, R. Susi, Sensitivity analysis in Gaussian Bayesian networks using a divergence measure, Communications in Statistics: Theory and Methods 36 (3) (2007) 523–539.
[10] M.A. Gómez-Villegas, P. Main, R. Susi, Extreme inaccuracies in Gaussian Bayesian networks, Journal of Multivariate Analysis 99 (2008) 1929–1940.
[11] M.A. Gómez-Villegas, P. Main, H. Navarro, R. Susi, Evaluating the difference between graph structures in Gaussian Bayesian networks, Expert Systems with Applications 38 (2011) 12409–12414.
[12] E.R. Hansen, R.R. Smith, Interval arithmetic in matrix computations, SIAM Journal of Numerical Analysis 4 (1994) 1–9.
[13] J. Horn, Positive definiteness and stability of interval matrices, SIAM Journal of Matrix Analysis and Applications 15 (1994) 175–184.
[14] J. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 2005.
[15] G.B. Huang, D.H. Wang, Y. Lan, Extreme learning machines: a survey, International Journal of Machine Learning and Cybernetics 2 (2) (2011) 107–122.
[16] F.V. Jensen, T. Nielsen, Bayesian Networks and Decision Graphs, sixth ed., Springer, New York, 2007.
[17] L.X. Jiang, H. Zhang, Z.H. Cai, A novel Bayes model: hidden Nadve Bayes, IEEE Transactions on Knowledge and Data Engineering 21 (10) (2009) 1361–1371.
[18] R.A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis, sixth ed., Prentice-Hall, 2007.
[19] U. Kjærulff, L.C. van der Gaag, Making sensitivity analysis computationally efficient, in: Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, San Francisco, CA, USA, 2000, pp. 315–325.
[20] S. Kullback, R.A. Leibler, On information and sufficiency, Annals of Mathematical Statistics 22 (1951) 79–86.
[21] K.B. Laskey, Sensitivity analysis for probability assessments in Bayesian networks, IEEE Transactions on Systems, Man and Cybernetics 25 (1995) 901–909.
[22] S.L. Lauritzen, Graphical Models, Clarendon Press, Oxford, 1996.
[23] P. Main, H. Navarro, Analyzing the effect of introducing a kurtosis parameter in Gaussian Bayesian networks, Reliability Engineering and System Safety 94 (2009) 922–926.
[24] R.E. McCulloch, Local model influence, Journal American Statistical Association 84 (1989) 473–478.
[25] S.L. Normand, D. Tritchler, Parameter updating in Bayes network, Journal American Statistical Association 87 (1992) 1109–1115.
[26] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, CA, 1988.
[27] R Development Core Team, 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>.
[28] S. Renooij. Bayesian network sensitivity to arc-removal, In: Proceedings of the Fifth European Workshop on Probabilistic Graphical Models, Helsinki, 2010.
[29] R. Shachter, C. Kenley, Gaussian influence diagrams, Management Science 35 (1989) 527–550.
[30] L.C. Van der Gaag, S. Renooij, V.M.H. Coup é, Sensitivity analysis of probabilistic networks, in: Advances in Probabilistic Graphical Models, Studies in Fuzziness and Soft Computing, vol. 213, Springer, Berlin, 2007, pp. 103–124.
[31] X.Z. Wang, C.R. Dong, Improving generalization of fuzzy if–then rules by maximizing fuzzy entropy, IEEE Transactions on Fuzzy Systems 17 (3) (2009) 556–567.
[32] X.Z. Wang, C.R. Dong, J.H. Yan, Maximum ambiguity based sample selection in fuzzy decision tree induction, IEEE Transactions on Knowledge and Data Engineering, 2011. 10.1109/TKDE.2011.67.
[33] J. Whittaker, Graphical Models in Applied Multivariate Statistics, Wiley, England, 1990.
[34] L. Xu, J. Pearl, Structuring causal tree models with continuous variables, Uncertainty in Artificial Intelligence 3 (1989) 209–222.