Reasoning with fuzzy and uncertain evidence using epistemic random fuzzy sets: general framework and practical models[1]

Thierry Denœux[a,b]

[a]*Université de technologie de Compiègne, CNRS*
UMR 7253 Heudiasyc, Compiègne, France
[b]*Institut universitaire de France, Paris, France*

Abstract

We introduce a general theory of epistemic random fuzzy sets for reasoning with fuzzy or crisp evidence. This framework generalizes both the Dempster-Shafer theory of belief functions, and possibility theory. Independent epistemic random fuzzy sets are combined by the generalized product-intersection rule, which extends both Dempster's rule for combining belief functions, and the product conjunctive combination of possibility distributions. We introduce Gaussian random fuzzy numbers and their multi-dimensional extensions, Gaussian random fuzzy vectors, as practical models for quantifying uncertainty about scalar or vector quantities. Closed-form expressions for the combination, projection and vacuous extension of Gaussian random fuzzy numbers and vectors are derived.

Keywords: Belief functions, evidence theory, possibility theory, random sets, uncertainty.

1. Introduction

The Dempster-Shafer (DS) theory of belief functions [29] and possibility theory [38] were introduced independently in the late 1970's as non-probabilistic frameworks for reasoning with uncertainty [11, 10]. The former approach is based on the idea of representing elementary pieces of evidence as completely monotone capacities, or *belief functions*, and combining them using an operator known as the product-intersection rule or Dempster's rule. As probability measures are special belief functions, and Dempster's rule extends Bayesian conditioning, DS can be seen as an extension of Bayesian probability theory, particularly suitable to reasoning with severe uncertainty. There is also a strong relation between DS theory and the theory of random sets [23]: specifically, any random set induces a belief function and, conversely, any belief function can be seen as being induced by some random set [26]. In DS theory, a random set underlying a belief function does not represent a random mechanism for generating sets of outcomes, but the imprecise meanings of a piece of evidence under

---

different interpretations with known probabilities [30]. To avoid confusion, we use the term *epistemic random set* for random sets representing evidence in DS theory.

In contrast, possibility theory originates from the theory of fuzzy sets [36]. In this approach, a fuzzy statement about the variable of interest, seen as a flexible constraint on its precise but unknown value in some domain Θ, induces a possibility measure and a dual necessity measure on Θ. Interestingly, a necessity measure is a belief function, and the dual possibility measure is the corresponding plausibility function, but the converse is not true (a belief function is not, in general, a necessity measure). For this reason, possibility theory has sometimes been presented as "a special branch of evidence theory" (another name for DS theory) [21, page 187]. However, combining two necessity measures by Dempster's rule yields a belief function that is no longer a necessity measure: this combination rule is, thus, not compatible with possibilistic reasoning. In contrast, possibility theory has its own conjunctive combination operators based on triangular norms (or t-norms) [16]. Possibility and DS theory are, thus, two distinct models of uncertain reasoning based on related knowledge representation languages but different information processing mechanisms.

In a companion paper [9], we have revisited Zadeh's notion of "evidence of the second kind", defined as a pair $(X, \Pi_{(Y|X)})$ in which X is a discrete random variable on a set Ω and $\Pi_{(Y|X)}$ a collection of conditional possibility distributions of a variable Y given $X = x$, for all $x \in \Omega$. If random variable X is constant, we get a unique possibility distribution for variable Y; if the conditional possibility distributions $\Pi_{(Y|X)}$ take values in $\{0, 1\}$, then the pair $(X, \Pi_{(Y|X)})$ defines a random set equivalent to a DS mass function. The mappings associating, to each event, its expected necessity and its expected possibility are, respectively, belief and plausibility functions. In this framework, a possibility distribution thus represents certain but fuzzy evidence, while a DS mass function is a model of uncertain and crisp evidence. In general, a pair $(X, \Pi_{(Y|X)})$ defines an *epistemic random fuzzy set*, allowing us to describe evidence that is both uncertain and fuzzy. (The term "epistemic" emphasizes the distinction between this interpretation and that of random fuzzy sets as mechanisms for generating fuzzy data considered, for instance in [28, 17]). In [9], we have proposed a family of combination rules for epistemic random fuzzy sets in the finite setting, generalizing both Dempster's rule and the conjunctive combination rules of possibility theory. One of these rules, based on the product t-norm, is associative and arguably well suited for combining independent evidence. Equipped with this combination rule (called here the *generalized product-intersection rule*), the theory of epistemic random fuzzy sets can be seen as an extension of both DS theory and possibility theory, making it possible to combine evidence of various types, including expert assessments (possibly expressed in natural language), sensor information, and statistical evidence about a model parameter.

In this paper, drawing from mathematical results presented by Couso and Sánchez in [2], we give a more general exposition of the theory of epistemic fuzzy sets, considering arbitrary probability and measurable spaces. We define combination, marginalization and vacuous extension operations of random fuzzy sets in this general setting, laying the foundations of a wide-ranging theory of uncertainty encompassing DS and possibility theories as special cases. Finally, for the important case where the frame of discernment is $\mathbb{R}^p$, we propose Gaussian random fuzzy numbers and vectors as a practical model, generalizing both Gaussian random

variables and vectors on the one hand, and Gaussian possibility distributions on the other hand.

The rest of this paper is organized as follows. Classical models (including random sets, fuzzy sets and possibility theory) are first recalled in Section 2. Epistemic random fuzzy sets are then introduced in a general setting in Section 3. Finally, Gaussian random fuzzy numbers and vectors are studied, respectively, in Sections 4 and 5, and Section 6 concludes the paper.

2. Classical models

In this section, we recall the main definitions and results pertaining to the two models of uncertainty generalized in this paper: random sets and belief functions on the one hand (Section 2.1), fuzzy sets and possibility theory on the other hand (Section 2.2).

2.1. Random sets and belief functions

Whereas belief functions in the finite setting can be introduced without any reference to random sets [29], the mathematical framework of random sets is useful to analyze belief functions in more general spaces, and to define the practical models needed, e.g., in statistical applications. Important references about the link between random sets and belief functions include [26] and [2].

Let $(\Omega, \sigma_\Omega, P)$ be a probability space, (Θ, σ_Θ) a measurable space, and $\overline{X}$ a mapping from Ω to 2^Θ. The *upper and lower inverses* of $\overline{X}$ are defined, respectively, as follows:

$$\overline{X}^*(B) = B^* = \{\omega \in \Omega : \overline{X}(\omega) \cap B \neq \emptyset\} \tag{1a}$$

$$\overline{X}_*(B) = B_* = \{\omega \in \Omega : \emptyset \neq \overline{X}(\omega) \subseteq B\} \tag{1b}$$

for all $B \subseteq \Theta$. It is easy to check that

$$B^* \cap (B^c)_* = \emptyset$$

and

$$B^* \cup (B^c)_* = \{\omega \in \Omega : \overline{X}(\omega) \neq \emptyset\} = \Theta^*,$$

where B^c denotes the complement of B in Θ.

The mapping $\overline{X}$ is said to be $\sigma_\Omega - \sigma_\Theta$ *strongly measurable* [26] if, for all $B \in \sigma_\Theta$, $B^* \in \sigma_\Omega$ (or, equivalently, if for all $B \in \sigma_\Theta$, $B_* \in \sigma_\Omega$). The tuple $(\Omega, \sigma_\Omega, P, \Theta, \sigma_\Theta, \overline{X})$ is called a *random set*. When there is no confusion about the domain and co-domain, we will call the $\sigma_\Omega - \sigma_\Theta$ strongly measurable mapping $\overline{X}$ itself a *random set*.

In the special case where $|\overline{X}(\omega)| = 1$ for all $\omega \in \Omega$, we can define the mapping $X : \Omega \to \Theta$ such that $\overline{X}(\omega) = \{X(\omega)\}$ for all $\omega \in \Omega$. We then have $B^* = B_* = X^{-1}(B)$ for all $B \subseteq \Theta$, and X is $\sigma_\Omega - \sigma_\Theta$ measurable. The notion of random set thus extends that of random variable.

Belief and plausibility functions. From now on, we will assume, for simplicity, that $P(\Theta^*) = 1$. (If not verified, this property can be enforced by conditioning P on Θ^*). Let P^* and P_* be the lower and upper probability measures associated with random set $\overline{X}$, defined as the mappings from σ_Θ to $[0, 1]$ such that

$$P_*(B) = P(B_*) \tag{2}$$

and

$$P^*(B) = P(B^*) = 1 - P_*(B^c), \tag{3}$$

for all $B \in \sigma_\Theta$. Mapping P_* is a completely monotone capacity, i.e., a *belief function,* and P^* is the dual *plausibility function* [26, Proposition 1]. In the following, they will be denoted, respectively, as $Bel_{\overline{X}}$ and $Pl_{\overline{X}}$. The corresponding *contour function* is defined as the mapping $pl_{\overline{X}}$ from Θ to $[0, 1]$ such that

$$pl_{\overline{X}}(\theta) = Pl_{\overline{X}}(\{\theta\})$$

for all $\theta \in \Theta$. The subsets $\overline{X}(\omega) \subseteq \Theta$, for all $\omega \in \Omega$, are called the *focal sets* of $Bel_{\overline{X}}$.

Interpretation. In DS theory, Ω represents a set of interpretations of a piece of evidence about a variable $\boldsymbol{\theta}$ taking values in set Θ (called the *frame of discernment*). If interpretation $\omega \in \Omega$ holds, we know that $\boldsymbol{\theta} \in \overline{X}(\omega)$, and nothing more. For any $A \in \sigma_\Omega$, $P(A)$ is the (subjective) probability that the true interpretation lies in A. For any $B \in \sigma_\Theta$, the degree of belief $Bel_{\overline{X}}(B)$ is then a measure of support of the proposition "$\boldsymbol{\theta} \in B$" given the evidence, while the degree of plausibility $Pl_{\overline{X}}(B)$ is a measure of lack of support for the proposition "$\boldsymbol{\theta} \notin B$". Under this interpretation, the random set $\overline{X}$ represents a state of knowledge: it can be said to be *epistemic.*

Vacuous random set. A constant random set $(\Omega, \sigma_\Omega, P, \Theta, \sigma_\Theta, \overline{X})$ such that $\overline{X}(\omega) = \Theta$ for all $\omega \in \Omega$ is said to be *vacuous.* For such a random set, we have $Bel_{\overline{X}}(A) = 0$ for all $A \in \sigma_\Theta \setminus \{\Theta\}$ and $Pl_{\overline{X}}(A) = 1$ for all $A \in \sigma_\Theta \setminus \{\emptyset\}$. A vacuous random set represents complete ignorance about $\boldsymbol{\theta}$.

Finite case. Assume that Θ is finite, and $\sigma_\Theta = 2^\Theta$. The *Möbius inverse* of $Bel_{\overline{X}}$ is the mapping $m_{\overline{X}}$ from 2^Θ to $[0,1]$ such that

$$m_{\overline{X}}(B) = \sum_{C \subseteq B} (-1)^{|B|-|C|} Bel_{\overline{X}}(C),$$

for all $B \subseteq \Theta$. It verifies $m(B) \geq 0$ for all $B \subseteq \Omega$, $\sum_{B \subseteq \Omega} m(B) = 1$ and $m(\emptyset) = 0$. The belief and plausibility can be computed from $m_{\overline{X}}$, respectively, as

$$Bel_{\overline{X}}(B) = \sum_{C \subseteq B} m_{\overline{X}}(C) \quad \text{and} \quad Pl_{\overline{X}}(B) = \sum_{C \cap B \neq \emptyset} m_{\overline{X}}(C),$$

for all $B \subseteq \Theta$.

Random closed intervals. Random closed intervals are particularly simple models allowing us to define belief functions on the real line [4, 34, 6]. Let $(\Omega, \sigma_\Omega, P)$ be a probability space and X, Y two random variables $\Omega \to \mathbb{R}$ such that $P(\{\omega \in \Omega : X(\omega) \le Y(\omega)\}) = 1$. Then, the mapping $\overline{X} : \Omega \to 2^\mathbb{R}$ defined by $\overline{X}(\omega) = [X(\omega), Y(\omega)]$ is $\sigma_\Omega - \beta_\mathbb{R}$ strongly measurable, where $\beta_\mathbb{R}$ is the Borel σ-algebra on $\mathbb{R}$ (see a formal proof in [22]). This mapping defines a *random closed interval*. For a random closed interval $\overline{X} = [X, Y]$, we have [4]

$$Bel_{\overline{X}}([x, y]) = P([X, Y] \subseteq [x, y]) = P(X \ge x; Y \le y) \tag{4a}$$

and

$$Pl_{\overline{X}}([x, y]) = P([X, Y] \cap [x, y] \ne \emptyset) = 1 - P(X > y) - P(Y < x), \tag{4b}$$

for all $(x, y) \in \mathbb{R}^2$ such that $x \le y$. In particular, by letting x tend to $-\infty$ in (4), we obtain the *lower and upper cumulative distribution functions (cdf's)* of $\overline{X}$ as

$$F_*(y) = Bel_{\overline{X}}((-\infty, y]) = P(Y \le y) = F_Y(y) \tag{5a}$$

and

$$F^*(y) = Pl_{\overline{X}}((-\infty, y]) = P(X \le y) = F_X(y). \tag{5b}$$

Lower and upper expectation. Let $\overline{X}$ be a random set from $(\Omega, \sigma_\Omega, P)$ to $(\mathbb{R}, \beta_\mathbb{R})$. Following Dempster [3], we can define its *lower and upper expectations*, respectively, as

$$\mathbb{E}_*(\overline{X}) = \int_{-\infty}^{+\infty} x \, dF^*(x)$$

and

$$\mathbb{E}^*(\overline{X}) = \int_{-\infty}^{+\infty} x \, dF_*(x),$$

where $F_*(x) = Bel_{\overline{X}}((-\infty, x])$ and $F^*(x) = Pl_{\overline{X}}((-\infty, x])$ are the lower and upper cdf's of $\overline{X}$. When $\overline{X}$ is a random closed interval $[X, Y]$, it follows from (5) that $\mathbb{E}_*(\overline{X}) = \mathbb{E}(X)$ and $\mathbb{E}^*(\overline{X}) = \mathbb{E}(Y)$.

Dempster's rule. Consider two pieces of evidence represented by random sets

$$(\Omega_1, \sigma_1, P_1, \Theta, \sigma_\Theta, \overline{X}_1) \quad \text{and} \quad (\Omega_2, \sigma_2, P_2, \Theta, \sigma_\Theta, \overline{X}_2),$$

and the mapping $\overline{X}_\cap$ from $\Omega_1 \times \Omega_2$ to 2^Θ defined by $\overline{X}_\cap(\omega_1, \omega_2) = \overline{X}_1(\omega_1) \cap \overline{X}_2(\omega_2)$. If interpretations $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ both hold, we know that $\boldsymbol{\theta} \in \overline{X}_\cap(\omega_1, \omega_2)$, provided that $\overline{X}_1(\omega_1) \cap \overline{X}_2(\omega_2) \ne \emptyset$. Assume that $\overline{X}_\cap$ is $(\sigma_1 \otimes \sigma_2) - \sigma_\Theta$ strongly measurable, where $\sigma_1 \otimes \sigma_2$ is the tensor product σ-algebra over the Cartesian product $\Omega_1 \times \Omega_2$. The two pieces of evidence are said to be *independent* if, for any $A \in \sigma_1 \otimes \sigma_2$, the probability that A contains the true interpretations of the two pieces of evidence is the conditional probability

$$P_{12}(A) = (P_1 \times P_2)(A \mid \Theta^*) = \frac{(P_1 \times P_2)(A \cap \Theta^*)}{(P_1 \times P_2)(\Theta^*)}, \tag{6}$$

where $P_1 \times P_2$ is the product measure satisfying $(P_1 \times P_2)(A_1 \times A_2) = P_1(A_1)P_2(A_2)$ for all $A_1 \in \sigma_1$, $A_2 \in \sigma_2$, and

$$\Theta^* = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \overline{X}_\cap(\omega_1, \omega_2) \neq \emptyset\}$$

is the set of noncontradictory pairs of interpretations. The quantity

$$\kappa = 1 - (P_1 \times P_2)(\Theta^*) = (P_1 \times P_2)(\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \overline{X}_\cap(\omega_1, \omega_2) = \emptyset\})$$

is called the *degree of conflict* between the two pieces of evidence. The combined random set

$$(\Omega_1 \times \Omega_2, \sigma_1 \otimes \sigma_2, P_{12}, \Theta, \sigma_\Theta, \overline{X}_\cap)$$

is called the *orthogonal sum* of the two pieces of evidence, and is denoted by $\overline{X}_1 \oplus \overline{X}_2$. This combination rule, first introduced by Dempster in [3], is called the *product-intersection rule*, or *Dempster's rule of combination*.

We can remark that Dempster's rule is usually viewed as an operation to combine belief functions, whereas it is defined here as an operation to combine *random sets*. This distinction is immaterial in the standard setting, as the orthogonal sum of two belief functions does not depend on their particular random set representations and can be defined without reference to the random set framework [32]. However, it becomes crucial when considering random fuzzy sets as a model for generating belief functions, as done in this paper. We will come back to this important point in Section 3.2.

Any vacuous random set is obviously a neutral element for Dempster's rule. The following important proposition states that pieces of evidence can be combined by Dempster's rule in any order.

Proposition 1. *Dempster's rule is commutative and associative.*

Proof. See Appendix A. □

Example 1. *Let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ be two independent normal random variables and consider the random intervals $\overline{X}_1 = [X_1, +\infty)$ and $\overline{X}_2 = (-\infty, X_2]$. The degree of conflict between $\overline{X}_1$ and $\overline{X}_2$ is*

$$\kappa = P(X_1 > X_2) = P(X_2 - X_1 < 0) = \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right),$$

where Φ is the standard normal cdf. The orthogonal sum of $\overline{X}_1$ and $\overline{X}_2$ is the random closed interval $[X_1', X_2']$, where (X_1', X_2') is the two-dimensional random vector with distribution equal the conditional distribution of (X_1, X_2) given $X_1 \leq X_2$. Its density is

$$f_{X_1', X_2'}(x_1, x_2) = \frac{\sigma_1^{-1}\sigma_2^{-1}\phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right)\phi\left(\frac{x_2 - \mu_2}{\sigma_2}\right)I(x_1 \leq x_2)}{\Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)},$$

where ϕ is the standard normal probability density function (pdf) and $I(\cdot)$ is the indicator function.

The following proposition states that the contour function of the orthogonal sum of two independent random sets $\overline{X}_1$ and $\overline{X}_2$ is proportional to the product of the contour functions of $\overline{X}_1$ and $\overline{X}_2$.

Proposition 2. *Let $\overline{X}_1$ and $\overline{X}_2$ be two independent random sets on the same domain Θ, with contour functions $pl_{\overline{X}_1}$ and $pl_{\overline{X}_2}$. For any $\theta \in \Theta$,*

$$pl_{\overline{X}_1 \oplus \overline{X}_2}(\theta) = \frac{pl_{\overline{X}_1}(\theta)pl_{\overline{X}_2}(\theta)}{1 - \kappa}, \tag{7}$$

where κ is the degree of conflict between $\overline{X}_1$ and $\overline{X}_2$.

Proof. We have

$$
\begin{aligned}
pl_{\overline{X}_1 \oplus \overline{X}_2}(\theta) &= \frac{(P_1 \times P_2)(\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \theta \in \overline{X}_\cap(\omega_1, \omega_2)\})}{1 - \kappa} \\
&= \frac{(P_1 \times P_2)(\{\omega_1 \in \Omega_1 : \theta \in \overline{X}_1(\omega_1)\} \times \{\omega_2 \in \Omega_2 : \theta \in \overline{X}_2(\omega_2)\})}{1 - \kappa} \\
&= \frac{P_1(\{\omega_1 \in \Omega_1 : \theta \in \overline{X}_1(\omega_1)\}) \cdot P_2(\{\omega_2 \in \Omega_2 : \theta \in \overline{X}_2(\omega_2)\})}{1 - \kappa} \\
&= \frac{pl_{\overline{X}_1}(\theta)pl_{\overline{X}_2}(\theta)}{1 - \kappa}.
\end{aligned}
$$

$\square$

Example 2. *Let us consider again the two random intervals of Example 1. The contour functions of $\overline{X}_1$ and $\overline{X}_2$ are, respectively,*

$$pl_{\overline{X}_1}(x) = P(X_1 \leq x) = \Phi\left(\frac{x - \mu_1}{\sigma_1}\right)$$

and

$$pl_{\overline{X}_2}(x) = P(X_2 \geq x) = 1 - \Phi\left(\frac{x - \mu_2}{\sigma_2}\right).$$

Now, the contour function of $\overline{X}_1 \oplus \overline{X}_2$ is

$$
\begin{aligned}
pl_{\overline{X}_1 \oplus \overline{X}_2}(x) &= P(X_1' \leq x \leq X_2') \\
&= \int_{-\infty}^{x} \int_{x}^{+\infty} f_{X_1', X_2'}(x_1, x_2) dx_2 dx_1 \\
&= \left[\Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\right]^{-1} \int_{-\infty}^{x} \int_{x}^{+\infty} \sigma_1^{-1}\sigma_2^{-1} \phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right) \phi\left(\frac{x_2 - \mu_2}{\sigma_2}\right) dx_2 dx_1 \\
&= \frac{\Phi\left(\frac{x - \mu_1}{\sigma_1}\right)\left[1 - \Phi\left(\frac{x - \mu_2}{\sigma_2}\right)\right]}{\Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)} = \frac{pl_{\overline{X}_1}(x)pl_{\overline{X}_2}(x)}{1 - \kappa}.
\end{aligned}
$$

Marginalization and vacuous extension. Let us now consider the case where we have two variables $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ with domains Θ_1 and Θ_2. (The case of n variables is not more difficult conceptually but it requires heavier notations). Let σ_{Θ_1} and σ_{Θ_2} be σ-algebras defined, respectively, on Θ_1 and Θ_2. Let $\Theta_{12} = \Theta_1 \times \Theta_2$ and $\sigma_{\Theta_{12}} = \sigma_{\Theta_1} \otimes \sigma_{\Theta_2}$. Let $\overline{X}_{12}$ be a random set from $(\Omega, \sigma_\Omega, P)$ to $(\Theta_{12}, \sigma_{\Theta_{12}})$, and $\overline{X}_1$ the mapping from Ω to 2^{Θ_1} that maps each $\omega \in \Omega$ to the *projection* of $\overline{X}_{12}(\omega)$ onto Θ_1:

$$\overline{X}_1(\omega) = \overline{X}_{12}(\omega) \downarrow \Theta_1 = \{\theta_1 \in \Theta_1 : \exists \theta_2 \in \Theta_2, (\theta_1, \theta_2) \in \overline{X}_{12}(\omega)\}.$$

It is easy to see that $\overline{X}_1$ is $\sigma_\Omega - \sigma_{\Theta_1}$ measurable: for any $B \in \sigma_{\Theta_1}$,

$$
\begin{aligned}
\overline{X}_1^*(B) &= \{\omega \in \Omega : \overline{X}_1(\omega) \cap B \neq \emptyset\} \\
&= \{\omega \in \Omega : \overline{X}_{12}(\omega) \cap (B \times \Theta_2) \neq \emptyset\} \\
&= \overline{X}_{12}^*(B \times \Theta_2).
\end{aligned}
$$

As $B \times \Theta_2 \in \sigma_{\Theta_{12}}$ and $\overline{X}_{12}$ is $\sigma_\Omega - \sigma_{\Theta_{12}}$ strongly measurable, it thus follows that $\overline{X}_1^*(B) \in \sigma_\Omega$. The random set $\overline{X}_1$ will be called the *marginal* of $\overline{X}_{12}$ on Θ_1.

Conversely, let $\overline{X}_1$ be a random set from (Ω, σ_Ω) to $(\Theta_1, \sigma_{\Theta_1})$ and let $\overline{X}_{1\uparrow 2}$ be the mapping from Ω to Θ_{12} defined by

$$\overline{X}_{1\uparrow(1,2)}(\omega) = \overline{X}_1(\omega) \times \Theta_2.$$

For any $B \in \sigma_{\Theta_{12}}$,

$$
\begin{aligned}
\overline{X}_{1\uparrow(1,2)}^*(B) &= \{\omega \in \Omega : \overline{X}_{1\uparrow 2}(\omega) \cap B \neq \emptyset\} \\
&= \{\omega \in \Omega : \overline{X}_1(\omega) \cap (B \downarrow \Theta_1) \neq \emptyset\} \\
&= \overline{X}_1^*(B \downarrow \Theta_1).
\end{aligned}
$$

If for all $B \in \sigma_{\Theta_{12}}$, $\overline{X}_1^*(B \downarrow \Theta_1) \in \sigma_\Omega$, then $\overline{X}_{1\uparrow(1,2)}$ is $\sigma_\Omega - \sigma_{\Theta_{12}}$ strongly measurable. It is said to be the *vacuous extension* of $\overline{X}_1$ in $\Theta_1 \times \Theta_2$.

We say that a random set $\overline{X}_{12}$ from $(\Omega, \sigma_\Omega, P)$ to $(\Theta_{12}, \sigma_{\Theta_{12}})$ with marginals $\overline{X}_1$ and $\overline{X}_2$ is *noninteractive* if it is equal to the orthogonal sum of its marginals, i.e.,

$$\overline{X}_{12} = \overline{X}_{1\uparrow(1,2)} \oplus \overline{X}_{2\uparrow(1,2)} \quad \text{denoted by} \quad \overline{X}_1 \oplus \overline{X}_2.$$

Example 3. *Let (X_1, X_2) be a two dimensional random vector from $(\Omega, \sigma_\Omega, P)$ to $(\mathbb{R}^2, \beta_{\mathbb{R}^2})$ and consider the mapping $\overline{X}_{12} : \Omega \to 2^{\mathbb{R}^2}$ defined as*

$$\overline{X}_{12}(\omega) = (-\infty, X_1(\omega)] \times (-\infty, X_2(\omega)].$$

This mapping defines a random set [23, page 3]. Its marginals are the random closed intervals $(-\infty, X_1]$ and $(-\infty, X_2]$. If X_1 and X_2 are independent, then $\overline{X}_{12} = (-\infty, X_1] \oplus (-\infty, X_2]$ and $\overline{X}_{12}$ is noninteractive.

2.2. Fuzzy sets and possibility theory

A *fuzzy subset* of a set Θ is a pair $\widetilde{F} = (\Theta, \mu_{\widetilde{F}})$, where $\mu_{\widetilde{F}}$ is a mapping from Θ to $[0, 1]$, called the *membership function* of $\widetilde{F}$. Each number $\mu_{\widetilde{F}}(\theta)$ is interpreted as a degree of membership of element θ to the fuzzy set $\widetilde{F}$. In the following, to simplify the notation, we will identify fuzzy sets to their membership functions and write $\widetilde{F}(\theta)$ for $\mu_{\widetilde{F}}(\theta)$. The height of fuzzy set $\widetilde{F}$ is defined as

$$\mathsf{hgt}(\widetilde{F}) = \sup_{\theta \in \Theta} \widetilde{F}(\theta).$$

If $\mathsf{hgt}(\widetilde{F}) = 1$, $\widetilde{F}$ is said to be *normal*. For any $\alpha \in [0, 1]$, the (weak) α-cut of $\widetilde{F}$ is the set

$$^{\alpha}\widetilde{F} = \{\theta \in \Theta : \widetilde{F}(\theta) \geq \alpha\}.$$

Possibility and necessity measures. Let $\boldsymbol{\theta}$ be a variable taking values in Θ. Assume that we receive a piece of evidence telling us that "$\boldsymbol{\theta}$ is $\widetilde{F}$", where $\widetilde{F}$ is a normal fuzzy subset of Θ. This evidence induces a *possibility measure* $\Pi_{\widetilde{F}}$ from 2^{Θ} to $[0, 1]$ defined by

$$\Pi_{\widetilde{F}}(B) = \sup_{\theta \in B} \widetilde{F}(\theta), \tag{8}$$

for all $B \subseteq \Theta$. The number $\Pi_{\widetilde{F}}(B)$ is interpreted as the degree of possibility that $\boldsymbol{\theta} \in B$, given that $\boldsymbol{\theta}$ is $\widetilde{F}$ [38]. The corresponding *possibility distribution* is the mapping from Θ to $[0, 1]$ defined by

$$\pi_{\widetilde{F}}(\theta) = \Pi_{\widetilde{F}}(\{\theta\}) = \widetilde{F}(\theta),$$

i.e., it is identical to the membership function $\widetilde{F}$. The dual *necessity measure* is defined as

$$N_{\widetilde{F}}(B) = 1 - \Pi_{\widetilde{F}}(B^c) = \inf_{\theta \notin B} \left[1 - \widetilde{F}(\theta) \right]. \tag{9}$$

It can easily be shown that mapping $N_{\widetilde{F}} : 2^{\Omega} \to [0, 1]$ is completely monotone, i.e., it is a belief function, and $\Pi_{\widetilde{F}}$ is the dual plausibility function [15]. These belief and plausibility functions are formally induced by the random set $([0, 1], \beta_{[0,1]}, \lambda, \Theta, 2^{\Theta}, \overline{X})$, where $\beta_{[0,1]}$ is the Borel σ-algebra on $[0, 1]$, λ is the uniform probability measure, and $\overline{X}$ is the mapping $[0, 1] \to 2^{\Theta}$ defined by $\overline{X}(\alpha) = {}^{\alpha}\widetilde{F}$. However, as we will see in Section 3.2, it is important, when combining evidence, to distinguish between possibility distributions induced by fuzzy sets, and consonant belief functions induced by random sets.

Conjunctive combination of possibility distributions. Assume that we receive two independent pieces of information telling us that "$\boldsymbol{\theta}$ is $\widetilde{F}$" and "$\boldsymbol{\theta}$ is $\widetilde{G}$", where $\widetilde{F}$ and $\widetilde{G}$ are two fuzzy subsets of Θ. The conjunctive combination of these two pieces of evidence requires some notion of intersection between fuzzy sets. As reviewed in [13], the intersection operation can be extended to fuzzy sets using triangular norms (or t-norms for short). Given a t-norm $\top$, the $\top$-intersection of two fuzzy subsets $\widetilde{F}$ and $\widetilde{G}$ of the same domain Θ can be defined as

$$(\widetilde{F} \cap_{\top} \widetilde{G})(\theta) = \widetilde{F}(\theta) \top \widetilde{G}(\theta)$$

for all $\theta \in \Theta$. The most common choices for $\top$ are the minimum and product t-norms, as originally proposed by Zadeh [36]; the corresponding operations are called, respectively, the *minimum* and *product* intersections. However, the intersection of two normal fuzzy sets is generally not normal. To obtain a normal fuzzy set, as needed for the definitions of possibility and necessity measures in (8)-(9), we define the normalized $\top$-intersection as

$$(\widetilde{F} \cap_\top^* \widetilde{G})(\theta) = \begin{cases} \dfrac{\widetilde{F}(\theta) \top \widetilde{G}(\theta)}{\mathsf{hgt}(\widetilde{F} \cap_\top \widetilde{G})} & \text{if } \mathsf{hgt}(\widetilde{F} \cap_\top \widetilde{G}) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The fuzzy set $\widetilde{F} \cap_\top^* \widetilde{G}$ is normal provided that $\mathsf{hgt}(\widetilde{F} \cap_\top \widetilde{G}) > 0$. In general, the normalized intersection $\cap_\top^*$ associated with a t-norm $\top$ is not associative. A notable exception is the case where $\top$ is the product t-norm: the normalized product intersection, denoted by $\odot$, is associative (see [16], and a simple proof in [9]). By abuse of notation, we can use the same symbol to denote the conjunctive combination of possibility measures and the normalized product intersection of fuzzy sets, and write

$$\Pi_{\widetilde{F}} \odot \Pi_{\widetilde{G}} = \Pi_{\widetilde{F} \odot \widetilde{G}}.$$

As noted by Dubois and Prade [16, page 352], product intersection has a reinforcement effect that is appropriate when the information sources are assumed to be independent. The choice of the normalized product intersection for combining possibility distributions makes possibility theory fit in the framework of valuation-based systems [33] and allows for possibilistic reasoning with a large number of variables. The normalized product intersection operator also has an interesting property with respect to Gaussian fuzzy numbers, as recalled in the next paragraph.

Gaussian fuzzy numbers. A *fuzzy number* (or fuzzy interval) can be defined as a normal and convex fuzzy subset of the real line. In particular, a *Gaussian fuzzy number (GFN)* is a normal fuzzy subset of $\mathbb{R}$ with membership function

$$\varphi(x; m, h) = \exp\left(-\frac{h}{2}(x - m)^2\right),$$

where $m \in \mathbb{R}$ is the mode and $h \in [0, +\infty]$ is the precision. Such a fuzzy number will be denoted by $\mathsf{GFN}(m, h)$. If $h = 0$, $\varphi(x; m, h) = 1$ for all $x \in \mathbb{R}$: $\mathsf{GFN}(m, 0)$ is then maximally imprecise and identical to the whole real line, whatever the value of m. If $h = +\infty$, $\varphi(x; m, h) = I(x = m)$, where $I(\cdot)$ is the indicator function; the fuzzy number $\mathsf{GFN}(m, +\infty)$ is then maximally precise and equivalent to the real number m.

It can easily be shown that the family of GFN's is closed under the normalized product intersection (see, e.g., [1]). More precisely, we have the following proposition, proved in [1].

Proposition 3. *For any $x \in \mathbb{R}$,*

$$\varphi(x; m_1, h_1) \cdot \varphi(x; m_2, h_2) = \exp\left(-\frac{h_1 h_2 (m_1 - m_2)^2}{2(h_1 + h_2)}\right) \varphi(x; m_{12}, h_{12}),$$

with

$$m_{12} = \frac{h_1 m_1 + h_2 m_2}{h_1 + h_2} \quad and \quad h_{12} = h_1 + h_2.$$

Consequently,

$$\mathsf{GFN}(m_1, h_1) \odot \mathsf{GFN}(m_2, h_2) = \mathsf{GFN}(m_{12}, h_{12}),$$

and

$$\mathsf{hgt}\left[\mathsf{GFN}(m_1, h_1) \cdot \mathsf{GFN}(m_2, h_2)\right] = \exp\left(-\frac{h_1 h_2 (m_1 - m_2)^2}{2(h_1 + h_2)}\right). \tag{10}$$

Marginalization and cylindrical extension. Let us now assume that we have two variables $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ jointly constrained by a possibility distribution $\pi_{\widetilde{F}}$, where $\widetilde{F}$ is a fuzzy subset of $\Theta_{12} = \Theta_1 \times \Theta_2$. As a result of (8), variable $\boldsymbol{\theta}_1$ alone is constrained by the possibility distribution

$$\pi_1(\theta_1) = \Pi(\{\theta_1\} \times \Theta_2) = \sup_{\theta_2 \in \Theta_2} \pi_{\widetilde{F}}(\theta_1, \theta_2) = \sup_{\theta_2 \in \Theta_2} \widetilde{F}(\theta_1, \theta_2) = (\widetilde{F} \downarrow \Theta_1)(\theta_1),$$

where $\widetilde{F} \downarrow \Theta_1$ is the *projection* of $\widetilde{F}$ on Θ_1. We say that π_1 is the *marginal* of $\pi_{\widetilde{F}}$ on Θ_1. Conversely, given a possibility distribution $\pi_{\widetilde{F}_1}$, where $\widetilde{F}_1$ is a fuzzy subset of Θ_1, its *cylindrical extension* in $\Theta_1 \times \Theta_2$ is the possibility distribution $\pi_{\widetilde{F}_1 \times \Theta_2}$ defined as

$$\pi_{\widetilde{F}_1 \times \Theta_2}(\theta_1, \theta_2) = \pi_{\widetilde{F}_1}(\theta_1)$$

for all $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$. We say that the joint possibility distribution $\pi_{\widetilde{F}}$ on Θ_{12} is *noninteractive* with respect to the product intersection if it is the product of its marginals:

$$\pi_{\widetilde{F}}(\theta_1, \theta_2) = \pi_{\widetilde{F} \downarrow \Theta_1}(\theta_1) \cdot \pi_{\widetilde{F} \downarrow \Theta_2}(\theta_2).$$

Example 4. *Let π_{12} be the possibility distribution on $\mathbb{R}^2$ defined by*

$$\pi_{12}(x_1, x_2) = \exp\left(-\frac{h_1}{2}(x_1 - m_1)^2 - \frac{h_2}{2}(x_2 - m_2)^2\right)$$

$$= \exp\left(-\frac{h_1}{2}(x_1 - m_1)^2\right) \exp\left(-\frac{h_2}{2}(x_2 - m_2)^2\right).$$

Its marginals are

$$\pi_1(x_1) = \max_{\theta_2} \pi_{12}(x_1, x_2) = \exp\left(-\frac{h_1}{2}(x_1 - m_1)^2\right)$$

and

$$\pi_2(x_2) = \max_{\theta_1} \pi_{12}(x_1, x_2) = \exp\left(-\frac{h_2}{2}(x_2 - m_2)^2\right).$$

Consequenty, π_{12} is noninteractive with respect to the product intersection.

11

3. Epistemic random fuzzy sets

The proposed epistemic random fuzzy set model is introduced in this section. The main definitions are first given in Section 3.1, and the generalized product-intersection rule is introduced in Section 3.2. Marginalization and vacuous extension are then addressed in Section 3.3, and an application to statistical inference is briefly discussed in Section 3.4.

3.1. General definitions

As before, let $(\Omega, \sigma_\Omega, P)$ be a probability space and let (Θ, σ_Θ) be a measurable space. Let $\widetilde{X}$ by a mapping from Ω to the set $[0,1]^\Theta$ of fuzzy subsets of Θ. For any $\alpha \in [0,1]$, let $^\alpha \widetilde{X}$ be the mapping from Ω to 2^Θ defined as

$$^\alpha \widetilde{X}(\omega) = {}^\alpha [\widetilde{X}(\omega)],$$

where $^\alpha [\widetilde{X}(\omega)]$ is the weak α-cut of $\widetilde{X}(\omega)$. If for any $\alpha \in [0,1]$, $^\alpha \widetilde{X}$ is $\sigma_\Omega - \sigma_\Theta$ strongly measurable, the tuple $(\Omega, \sigma_\Omega, P, \Theta, \sigma_\Theta, \widetilde{X})$ is said to be a *random fuzzy set* (also called a *fuzzy random variable*) [2]. It is clear that the class of random fuzzy sets includes that of random sets, just as the class of fuzzy sets includes that of classical (crisp) sets.

Example 5. *Let M be a Gaussian random variable from $(\Omega, \sigma_\Omega, P)$ to $(\mathbb{R}, \beta_\mathbb{R})$, with mean μ and standard deviation σ, and let $\widetilde{X}$ be the mapping from Ω to $[0,1]^\mathbb{R}$ that maps each $\omega \in \Omega$ to the triangular fuzzy number with mode $M(\omega)$ and support $[M(\omega) - a, M(\omega) + a]$:*

$$\widetilde{X}(\omega)(x) = \begin{cases} \frac{a - |x - M(\omega)|}{a} & \text{if } |x - M(\omega)| \leq a \\ 0 & \text{otherwise.} \end{cases}$$

for some $a > 0$. For any $\alpha \in [0,1]$, the α-cut of $\widetilde{X}(\omega)$ is

$$^\alpha \widetilde{X}(\omega) = [M(\omega) - a(1-\alpha), M(\omega) + a(1-\alpha)].$$

The random set $^\alpha \widetilde{X} : \omega \to {}^\alpha \widetilde{X}(\omega)$ is $\sigma_\Omega - \beta_\mathbb{R}$ strongly measurable (it is a random closed interval). Consequently, $\widetilde{X}$ is a random fuzzy set. In the following, such random fuzzy sets with domain $[0,1]^\mathbb{R}$ will be called random fuzzy numbers.

Interpretation. Here, as in [9], we use random fuzzy sets as a model of unreliable and fuzzy evidence. In this model, we see Ω as a *set of interpretations* of a piece of evidence about a variable $\boldsymbol{\theta}$ taking values in Θ. If interpretation $\omega \in \Omega$ holds, we know that "$\boldsymbol{\theta}$ is $\widetilde{X}(\omega)$", i.e., $\boldsymbol{\theta}$ is constrained by the possibility distribution $\pi_{\widetilde{X}(\omega)}$. We qualify such random fuzzy sets as *epistemic*, because they encode a state of knowledge about some variable $\boldsymbol{\theta}$. It should be noted that this semantics of random fuzzy sets is different from those reviewed in [2]. The conditional possibility interpretation developed in [2] is the closest to ours, since we also see the fuzzy sets $\widetilde{X}(\omega)$ as defining conditional possibility measures. However, in [2], the authors use the random fuzzy set formalism to model a situation in which we have two random experiments, one of which is completely determined; the family of possibility

distributions $\{\pi_{\widetilde{X}(\omega)} : \omega \in \Omega\}$ then models our knowledge about the relationship between the outcomes ω of the first experiment and the possible outcomes of the second one. This formalism allows the authors of [2] to compute lower and upper bounds on the probability of any event related to the second experiment. In contrast, our model does not rely on the notion of random experiment. In particular, we do not postulate the existence of an objective probability measure on Θ, and the belief and plausibility measures introduced below are not interpreted as lower and upper bounds on "true" probabilities.

Belief and plausibility. We say that random fuzzy set $\widetilde{X}$ is *normalized* if it verifies the following conditions:

1. For all $\omega \in \Omega$, $\widetilde{X}(\omega)$ is either the empty set, or a normal fuzzy set, i.e., $\mathsf{hgt}(\widetilde{X}(\omega)) \in \{0,1\}$.
2. $P(\{\omega \in \Omega : \widetilde{X}(\omega) = \emptyset\}) = 0$.

These conditions will be assumed in the rest of this section. For any $\omega \in \Omega$, let $\Pi_{\widetilde{X}}(\cdot \mid \omega)$ be the possibility measure on Θ induced by $\widetilde{X}(\omega)$:

$$\Pi_{\widetilde{X}}(B \mid \omega) = \sup_{\theta \in B} \widetilde{X}(\omega)(\theta), \tag{11}$$

and let $N_{\widetilde{X}}(\cdot \mid \omega)$ be the dual necessity measure:

$$N_{\widetilde{X}}(B \mid \omega) = \begin{cases} 1 - \Pi_{\widetilde{X}}(B^c \mid \omega) & \text{if } \widetilde{X}(\omega) \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Let $Bel_{\widetilde{X}}$ and $Pl_{\widetilde{X}}$ be the mappings from σ_Θ to $[0,1]$ defined as

$$Bel_{\widetilde{X}}(B) = \int_\Omega N(B \mid \omega) dP(\omega) \tag{12}$$

and

$$Pl_{\widetilde{X}}(B) = \int_\Omega \Pi(B \mid \omega) dP(\omega). \tag{13}$$

Function $Bel_{\widetilde{X}}$ is a belief function, and $Pl_{\widetilde{X}}$ is the dual plausibility function. As shown in [2, Lemma 6.2], they are induced by the random set $(\Omega \times [0,1], \sigma_\Omega \otimes \beta_{[0,1]}, P \otimes \lambda, \Theta, \sigma_\Theta, \overline{X})$, where $\overline{X} : \Omega \times [0,1] \to 2^\Theta$ is the multi-valued mapping defined as

$$\overline{X}(\omega, \alpha) = {}^\alpha \widetilde{X}(\omega). \tag{14}$$

As a consequence, $Bel_{\widetilde{X}}(B)$ and $Pl_{\widetilde{X}}(B)$ can also be written as follows:

$$Bel_{\widetilde{X}}(B) = \int_0^1 Bel_{\alpha \widetilde{X}}(B) d\alpha \tag{15a}$$

and

$$Pl_{\widetilde{X}}(B) = \int_0^1 Pl_{\alpha \widetilde{X}}(B) d\alpha. \tag{15b}$$

Lower and upper expectations of a random fuzzy number. Let $\widetilde{X}$ be a random fuzzy number (i.e., a random fuzzy set with domain $[0,1]^{\mathbb{R}}$), and let $\overline{X}$ be the corresponding random set defined by (14). We define the lower and upper expectations of $\widetilde{X}$ as the lower and upper expectations of $\overline{X}$, i.e., $\mathbb{E}_*(\widetilde{X}) = \mathbb{E}_*(\overline{X})$ and $\mathbb{E}^*(\widetilde{X}) = \mathbb{E}^*(\overline{X})$. It follows from (15) that

$$\mathbb{E}_*(\widetilde{X}) = \int_0^1 \mathbb{E}_*(^\alpha\widetilde{X})d\alpha \quad \text{and} \quad \mathbb{E}^*(\widetilde{X}) = \int_0^1 \mathbb{E}^*(^\alpha\widetilde{X})d\alpha. \tag{16}$$

Example 6. *Let us consider again the random fuzzy number of Example 5. Its lower and upper cdf's are, respectively, the mappings $x \to Bel_{\widetilde{X}}((-\infty, x])$ and $x \to Pl_{\widetilde{X}}((-\infty, x])$. Let us illustrate the calculation of the upper cdf first, using two methods.*

Method 1. From (11),

$$\Pi\left((-\infty, x] \mid \omega\right) = \sup_{x' \leq x} \widetilde{X}(\omega)(x') = \begin{cases} 1 & \text{if } M(\omega) \leq x \\ \frac{x - M(\omega) + a}{a} & \text{if } x < M(\omega) \leq x + a \\ 0 & \text{otherwise.} \end{cases}$$

Using (13), we get

$$Pl_{\widetilde{X}}((-\infty, x]) = P(M \leq x) \times 1 + P(x < M \leq x + a)\mathbb{E}\left[\frac{x - M + a}{a} \mid x < M \leq x + a\right]$$

$$= \Phi\left(\frac{x - \mu}{\sigma}\right) + \left[\Phi\left(\frac{x + a - \mu}{\sigma}\right) - \Phi\left(\frac{x - \mu}{\sigma}\right)\right] \times$$

$$\left(\frac{x + a}{a} - \mathbb{E}\left[M \mid x < M \leq x + a\right]\right).$$

Now, using a well-known result about the truncated normal distribution,

$$\mathbb{E}\left[M \mid x < M \leq x + a\right] = \mu + \sigma\frac{\phi\left(\frac{x - \mu}{\sigma}\right) - \phi\left(\frac{x + a - \mu}{\sigma}\right)}{\Phi\left(\frac{x + a - \mu}{\sigma}\right) - \Phi\left(\frac{x - \mu}{\sigma}\right)}.$$

After rearranging the terms, we finally obtain

$$Pl_{\widetilde{X}}((-\infty, x]) = \left(\frac{x + a - \mu}{a}\right)\Phi\left(\frac{x + a - \mu}{\sigma}\right) - \left(\frac{x - \mu}{a}\right)\Phi\left(\frac{x - \mu}{\sigma}\right) +$$

$$\frac{\sigma}{a}\left[\phi\left(\frac{x + a - \mu}{\sigma}\right) - \phi\left(\frac{x - \mu}{\sigma}\right)\right]. \tag{17}$$

Method 2. Let us now use (15b). We have

$$Pl_{\widetilde{X}}((-\infty, x]) = \int_0^1 P(M - a(1 - \alpha) \leq x)d\alpha$$

$$= \int_0^1 \Phi\left(\frac{x + a(1 - \alpha) - \mu}{\sigma}\right)d\alpha.$$
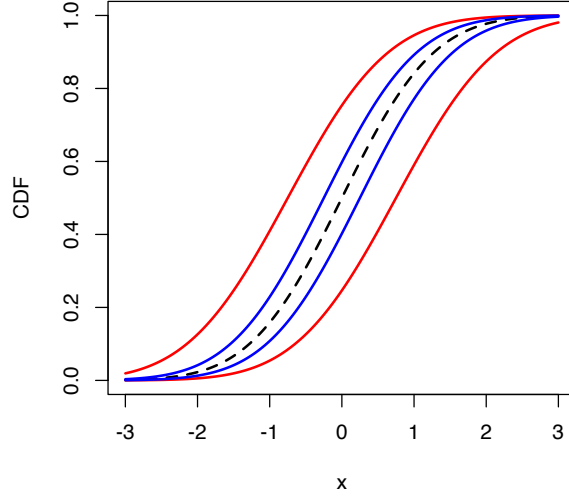
Figure 1: Lower and upper cdf's for the random fuzzy numbers studied in Examples 5 and 6, with $\mu = 0$, $\sigma = 1$, and $a = 0.5$ (blue curves) or $a = 1.5$ (red curves). The Gaussian cdf corresponding to $a = 0$ is shown as a broken line.

Using the formula

$$\int \Phi(u + vx)dx = \frac{1}{v}\left[(u + vx)\Phi(u + vx) + \phi(u + vx)\right] + C,$$

we get the same result as (17). *Using any of the two methods demonstrated above, we obtain the following expression for the lower cdf:*

$$Bel_{\widetilde{X}}((-\infty, x]) = \left(\frac{x - \mu}{a}\right)\Phi\left(\frac{x - \mu}{\sigma}\right) - \left(\frac{x - a - \mu}{a}\right)\Phi\left(\frac{x - a - \mu}{\sigma}\right) +$$
$$\frac{\sigma}{a}\left[\phi\left(\frac{x - \mu}{\sigma}\right) - \phi\left(\frac{x - a - \mu}{\sigma}\right)\right]. \quad (18)$$

It can easily be checked that, when $a = 0$,

$$Bel_{\widetilde{X}}((-\infty, x]) = Pl_{\widetilde{X}}((-\infty, x]) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Examples of functions $Bel_{\widetilde{X}}((-\infty, x])$ and $Pl_{\widetilde{X}}((-\infty, x])$ for different values of a are shown in Figure 1.

Now, the lower and upper expectations of $\widetilde{X}$ can be computed from (16) *as*

$$\mathbb{E}_*(\widetilde{X}) = \int_0^1 \mathbb{E}_*(^\alpha\widetilde{X})d\alpha = \int_0^1 [\mu - a(1 - \alpha)]d\alpha = \mu - \frac{a}{2},$$

and

$$\mathbb{E}^*(\widetilde{X}) = \int_0^1 \mathbb{E}^*(^\alpha\widetilde{X})d\alpha = \int_0^1 [\mu + a(1 - \alpha)]d\alpha = \mu + \frac{a}{2}.$$

15

3.2. Generalized product-intersection rule

Dempster's rule and the possibilistic product intersection rule recalled, respectively, in Sections 2.1 and 2.2 can be generalized to combine epistemic random fuzzy sets. Consider two epistemic random fuzzy sets $(\Omega_1, \sigma_1, P_1, \Theta, \sigma_\Theta, \widetilde{X}_1)$ and $(\Omega_2, \sigma_2, P_2, \Theta, \sigma_\Theta, \widetilde{X}_2)$ encoding independent pieces of evidence. The independence assumption means here that the relevant probability measure on the joint measurable space $(\Omega_1 \times \Omega_2, \sigma_1 \otimes \sigma_2)$ is the product measure $P_1 \times P_2$.

If interpretations $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ both hold, we know that "$\boldsymbol{\theta}$ is $\widetilde{X}_1(\omega_1)$" and "$\boldsymbol{\theta}$ is $\widetilde{X}_2(\omega_2)$". It is then natural to combine the fuzzy sets $\widetilde{X}_1(\omega_1)$ and $\widetilde{X}_2(\omega_2)$ by an intersection operator. As discussed in Section 2.2, normalized product intersection is a good candidate as it suitable for combining fuzzy information from independent sources and it is associative. We will thus consider the mapping $\widetilde{X}_\odot(\omega_1, \omega_2) = \widetilde{X}_1(\omega_1) \odot \widetilde{X}_2(\omega_2)$, which we will assume to be $\sigma_1 \otimes \sigma_2$-$\sigma_\Theta$ strongly measurable.

As in the crisp case recalled in Section 2.1, if $\mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)) = 0$, the two interpretations ω_1 and ω_2 are inconsistent and they must be discarded. If $\mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)) = 1$, the two interpretations are fully consistent. If $0 < \mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)) < 1$, ω_1 and ω_2 are *partially consistent*. As proposed in [9], instead of simply discarding only fully inconsistent pairs (ω_1, ω_2), it makes sense to give all pairs (ω_1, ω_2) a weight proportional to the *degree of consistency* between $\widetilde{X}_1(\omega_1)$ and $\widetilde{X}_2(\omega_2)$. This can be achieved by conditioning $P_1 \times P_2$ on the *fuzzy set* $\widetilde{\Theta}^*$ of consistent pairs of interpretations, with membership function

$$\widetilde{\Theta}^*(\omega_1, \omega_2) = \mathsf{hgt}\left(\widetilde{X}_1(\omega_1) \cdot \widetilde{X}_2(\omega_2)\right).$$

Using Zadeh's definition of a fuzzy event [37], we get the following expression for the conditional probability measure $\widetilde{P}_{12} = (P_1 \times P_2)(\cdot \mid \widetilde{\Theta}^*)$, for any $B \in \sigma_1 \otimes \sigma_2$:

$$\widetilde{P}_{12}(B) = \frac{(P_1 \times P_2)(B \cap \widetilde{\Theta}^*)}{(P_1 \times P_2)(\widetilde{\Theta}^*)} = \frac{\int_{\Omega_1} \int_{\Omega_2} B(\omega_1, \omega_2) \mathsf{hgt}\left(\widetilde{X}_1(\omega_1) \cdot \widetilde{X}_2(\omega_2)\right) dP_2(\omega_2) dP_1(\omega_1)}{\int_{\Omega_1} \int_{\Omega_2} \mathsf{hgt}\left(\widetilde{X}_1(\omega_1) \cdot \widetilde{X}_2(\omega_2)\right) dP_2(\omega_2) dP_1(\omega_1)},$$

where $B(\cdot, \cdot)$ denotes the indicator function of B. This conditioning operation, called *soft normalization* was first proposed in [35] in the finite case and with a different justification.

The combined random fuzzy set

$$(\Omega_1 \times \Omega_2, \sigma_1 \otimes \sigma_2, \widetilde{P}_{12}, \Theta, \sigma_\Theta, \widetilde{X}_\odot)$$

is called the *orthogonal sum* of the two pieces of evidence. This operation generalizes both Dempster's rule and the normalized product of possibility distribution. We will refer to it as the *generalized product-intersection rule*, and it will be denoted by the same symbol $\oplus$ as Dempster's rule. It is clear that $\widetilde{X} \oplus \overline{X}_0 = \widetilde{X}$ for any random fuzzy set $\widetilde{X}$ and any vacuous random set $\overline{X}_0$ on the same domain Θ. The degree of conflict between two random fuzzy sets $\widetilde{X}_1$ and $\widetilde{X}_2$ is naturally defined as

$$\kappa = 1 - (P_1 \times P_2)(\widetilde{\Theta}^*) = 1 - \int_{\Omega_1} \int_{\Omega_2} \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\right) dP_2(\omega_2) dP_1(\omega_1). \qquad (19)$$

The associativity of $\oplus$ was proved in [9] in the finite case; we give a similar proof in the general case.

Proposition 4. *The generalized product-intersection rule $\oplus$ for random fuzzy sets is commutative and associative.*

Proof. See Appendix B $\qquad\qquad\square$

The following proposition states that a counterpart of Proposition 2 is still valid when combining independent random fuzzy sets, i.e., the combined contour function is still proportional to the product of the contour functions.

Proposition 5. *Let $\widetilde{X}_1$ and $\widetilde{X}_2$ be two random fuzzy sets on the same domain Θ, with contour functions $pl_{\widetilde{X}_1}$ and $pl_{\widetilde{X}_2}$ and with degree of conflict κ defined by (19). The contour function $pl_{\widetilde{X}_1 \oplus \widetilde{X}_2}$ of $\widetilde{X}_1 \oplus \widetilde{X}_2$ verifies*

$$(pl_{\widetilde{X}_1 \oplus \widetilde{X}_2})(\theta) = \frac{pl_{\widetilde{X}_1}(\theta)pl_{\widetilde{X}_2}(\theta)}{1-\kappa}, \tag{20}$$

for all $\theta \in \Theta$.

Proof. We have

$$
\begin{aligned}
(pl_{\widetilde{X}_1 \oplus \widetilde{X}_2})(\theta) &= \frac{\int_{\Omega_1}\int_{\Omega_2} \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\cdot\widetilde{X}_2(\omega_2)\right)\widetilde{X}_\odot(\omega_1,\omega_2)(\theta)dP_2(\omega_2)dP_1(\omega_1)}{1-\kappa} \\
&= \frac{\int_{\Omega_1}\int_{\Omega_2} \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\cdot\widetilde{X}_2(\omega_2)\right)\frac{\widetilde{X}_1(\omega_1)(\theta)\widetilde{X}_2(\omega_2)(\theta)}{\mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\right)}dP_2(\omega_2)dP_1(\omega_1)}{1-\kappa} \\
&= \frac{\left(\int_{\Omega_1}\widetilde{X}_1(\omega_1)(\theta)dP_1(\omega_1)\right)\left(\int_{\Omega_2}\widetilde{X}_2(\omega_2)(\theta)dP_2(\omega_2)\right)}{1-\kappa} \\
&= \frac{pl_{\widetilde{X}_1}(\theta)pl_{\widetilde{X}_2}(\theta)}{1-\kappa}.
\end{aligned}
$$

$\qquad\qquad\square$

As remarked in Section 2.2, a belief function induced by a random fuzzy set is also induced by a random (crisp) set. However, combining random fuzzy sets or random crisp sets does not result in the same belief function in general. In particular, it is well-known that Dempster's rule does not preserve consonance. To combine two belief functions, we must, therefore, examine the evidence on which they are based, not only to determine whether the bodies of evidence are independent or not, but also to determine whether the evidence is fuzzy or crisp. This point is illustrated by the following example.
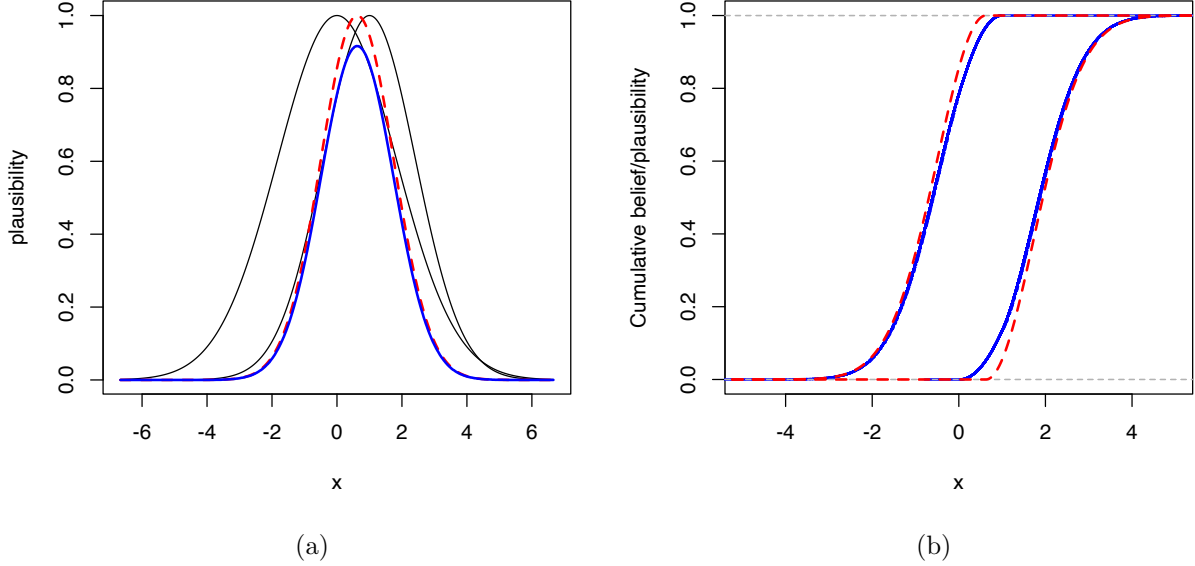
Figure 2: (a): Two Gaussian possibility distributions (black solid curves) with their normalized product intersection (red broken curve) and the contour function of the combined random set (blue solid curve). (b): Lower and upper cdf's of the combined possibility distribution (red broken curves) and of the combined random set (blue solid curves).

Example 7. *Consider the following two mappings from $\mathbb{R}$ to $[0, 1]$ represented in Figure 2a:*

$$\pi_1(x) = \mathsf{GFN}(0, 0.3), \quad \pi_2(x) = \mathsf{GFN}(1, 0.5).$$

If these two mappings are possibility distributions encoding fully reliable but fuzzy evidence, they correspond to "constant random fuzzy sets", i.e., mappings $\widetilde{X}_1(\omega) = \pi_1$ and $\widetilde{X}_2(\omega) = \pi_2$ with $P(\{\omega\}) = 1$. The combined random fuzzy set $\widetilde{X}_1 \oplus \widetilde{X}_2$ is then defined by $(\widetilde{X}_1 \oplus \widetilde{X}_2)(\omega) = \pi_1 \odot \pi_2$. From Proposition 3, the normalized product of two GFN's is a GFN. Here, we get the combined possibility distribution plotted as a red broken curve in Figure 2a:

$$(\pi_1 \odot \pi_2)(x) = \mathsf{GFN}(0.625, 0.8).$$

The corresponding lower and upper cumulative distribution functions (cdf's) are, respectively

$$Bel_{\widetilde{X}_1 \oplus \widetilde{X}_2}((-\infty, x]) = \begin{cases} 0 & \text{if } x \leq 0.625 \\ 1 - \exp\left(-0.4(x - 0.625)^2\right) & \text{if } x > 0.625 \end{cases}$$

and

$$Pl_{\widetilde{X}_1 \oplus \widetilde{X}_2}((-\infty, x]) = \begin{cases} \exp\left(-0.4(x - 0.625)^2\right) & \text{if } x \leq 0.625 \\ 1 & \text{if } x > 0.625. \end{cases}$$

18

These two functions are plotted as red broken curves in Figure 2b. Alternatively, as explained in Section 2.1, we may see π_1 and π_2 as encoding crisp but partially reliable evidence, in which case they define two independent consonant random intervals $\overline{X}_1(\alpha_1) = {}^{\alpha_1}\pi_1$ and $\overline{X}(\alpha_2) = {}^{\alpha_2}\pi_2$, where (α_1, α_2) has a uniform distribution on $[0,1]^2$. These two random intervals can be combined numerically using Monte-Carlo simulation, as explained in [20]. The contour function and the lower and upper cdf's are plotted as solid blue lines in Figures 2a and 2b, respectively. We notice that the contour functions are proportional, as a consequence of Proposition 2.

3.3. Marginalization and vacuous extension

Let us now consider again the case where we have two variables $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ with respective domains Θ_1 and Θ_2. Let $\widetilde{X}_{12}$ be a random fuzzy set from a probability space $(\Omega, \sigma_\Omega, P)$ to the measurable space $(\Theta_{12}, \sigma_{\Theta_{12}})$ with $\Theta_{12} = \Theta_1 \times \Theta_2$ and $\sigma_{\Theta_{12}} = \sigma_{\Theta_1} \otimes \sigma_{\Theta_2}$, where σ_{Θ_1} and σ_{Θ_2} are σ-algebras on Θ_1 and Θ_2, respectively. Let $\widetilde{X}_1$ be the mapping from Ω to $[0,1]^{\Theta_1}$ defined by

$$\widetilde{X}_1(\omega) = \widetilde{X}_{12}(\omega) \downarrow \Theta_1,$$

where, as before, $\downarrow$ denotes fuzzy set projection. If, for all $\alpha \in [0,1]$, the mapping ${}^{\alpha}\widetilde{X}_1$ is $\sigma_\Omega - \sigma_{\Theta_1}$ strongly measurable, then the random fuzzy set $\widetilde{X}_1$ is called the *marginal* of $\widetilde{X}_{12}$ on Θ_1.

Conversely, given a random fuzzy set $\widetilde{X}_1$ from $(\Omega, \sigma_\Omega, P)$ to $(\Theta_1, \sigma_{\Theta_1})$, let $\widetilde{X}_{1\uparrow(1,2)}$ be the mapping from Ω to $[0,1]^{\Theta_{12}}$ that maps each $\omega \in \Omega$ to the cylindrical extension of $\widetilde{X}_1(\omega)$ in Θ_{12}

$$\widetilde{X}_{1\uparrow(1,2)}(\omega) = \widetilde{X}_1(\omega) \times \Theta_2,$$

i.e., for all $(\theta_1, \theta_2) \in \Theta_{12}$,

$$\widetilde{X}_{1\uparrow(1,2)}(\omega)(\theta_1, \theta_2) = \widetilde{X}_1(\omega)(\theta_1).$$

If the mapping $\widetilde{X}_{1\uparrow(1,2)}$ is $\sigma_\Omega - \sigma_{\Theta_{12}}$ strongly measurable, then the random fuzzy set $\widetilde{X}_{1\uparrow(1,2)}$ is called the *vacuous extension* of $\widetilde{X}_1$ in Θ_{12}.

We say that a joint random fuzzy set is *noninteractive* if it is equal to the orthogonal sum of the vacuous extensions of its projections:

$$\widetilde{X}_{12} = \widetilde{X}_{1\uparrow(1,2)} \oplus \widetilde{X}_{2\uparrow(1,2)} \quad \text{denoted as} \quad \widetilde{X}_1 \oplus \widetilde{X}_2.$$

A particular kind of noninteractive random fuzzy sets will be studied in Section 5.3.

3.4. Application to statistical inference

Epistemic random fuzzy sets naturally arise in the context of statistical inference. As proposed by Shafer [29] and formally justified in [7][8], the information conveyed by the likelihood function in statistical inference problems can be represented by a consonant belief function, whose contour function is equal to the relative likelihood function. For a statistical model $f(\boldsymbol{x}, \theta)$, where $\boldsymbol{x} \in \mathcal{X}$ is the observation and $\theta \in \Theta$ is the unknown parameter, the

likelihood-based belief function $Bel(\cdot, \boldsymbol{x})$ on Θ after observing $\boldsymbol{x}$ is, thus, consonant and defined by the contour function

$$pl(\theta; \boldsymbol{x}) = \frac{L(\theta; \boldsymbol{x})}{\sup_{\theta' \in \Theta} L(\theta'; \boldsymbol{x})}, \tag{21}$$

where $L(\cdot, \boldsymbol{x}) : \theta \to f(\boldsymbol{x}; \theta)$ is the likelihood function, and it is assumed that the denominator in (21) is finite. The corresponding plausibility function is, thus, defined by

$$Pl(A; \boldsymbol{x}) = \sup_{\theta \in A} pl(\theta; \boldsymbol{x})$$

for any $A \subset \Theta$, i.e., it is a possibility measure. However, as noticed by Shafer in [29] and [31], and also discussed in [7], this construction is not compatible with Dempster's rule: if we consider two independent observations $\boldsymbol{x}$ and $\boldsymbol{x}'$, the belief function $Bel(\cdot; \boldsymbol{x}, \boldsymbol{x}')$ is not equal to the orthogonal sum $Bel(\cdot; \boldsymbol{x}) \oplus Bel(\cdot; \boldsymbol{x}')$, which is not consonant. As argued in [9], this problem disappears if we do not consider the likelihood-based belief function to be induced by a consonant random crisp set, but by a constant random fuzzy set $\widetilde{\theta}_{\boldsymbol{x}}$ with membership function $\widetilde{\theta}_{\boldsymbol{x}}(\theta) = pl(\theta; \boldsymbol{x})$. We can interpret $\widetilde{\theta}_{\boldsymbol{x}}$ as the fuzzy set of likely values of θ after observing $\boldsymbol{x}$. Combining the contour functions (21) by the normalized product intersection rule then yields the correct result, i.e., the constant random fuzzy set $\widetilde{\theta}_{\boldsymbol{x}, \boldsymbol{x}'}$ with membership function $\widetilde{\theta}_{\boldsymbol{x}, \boldsymbol{x}'}(\theta) = \widetilde{\theta}_{\boldsymbol{x}}(\theta) \odot \widetilde{\theta}_{\boldsymbol{x}'}(\theta)$.

Now, consider a prediction problem, where we want to predict the value of a random variable Y whose distribution also depends on θ. We can always write Y in the form $Y = \varphi(\theta, U)$, where U is a pivotal random variable with known distribution [19, 20]. After observing the data $\boldsymbol{x}$, our knowledge of θ is represented by the fuzzy set $\widetilde{\theta}_{\boldsymbol{x}}$. Conditionally on $U = u$, our knowledge of Y is, thus, represented by the fuzzy set $\widetilde{Y}(u) = \varphi(\widetilde{\theta}_{\boldsymbol{x}}, u)$, with membership function

$$\widetilde{Y}(u)(y) = \sup_{\theta : \varphi(\theta, u) = y} \widetilde{\theta}_{\boldsymbol{x}}(\theta).$$

The mapping $\widetilde{Y} : u \to \widetilde{Y}(u)$ is, then, a random fuzzy set representing statistical evidence about Y.

Example 8. *Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be an independent and identically distributed (iid) Gaussian sample with parent distribution $N(\theta, 1)$, and let $Y \sim N(\theta, 1)$. After observing a realization $\boldsymbol{x}$ of $\boldsymbol{X}$, the likelihood function is*

$$L(\theta; \boldsymbol{x}) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (x_i - \theta)^2\right).$$

Denoting by $\widehat{\theta}$ the sample mean, the fuzzy set $\widetilde{\theta}_{\boldsymbol{x}}$ of likely values of θ after observing $\boldsymbol{x}$ is the relative likelihood

$$\widetilde{\theta}_{\boldsymbol{x}}(\theta) = \frac{L(\theta; \boldsymbol{x})}{L(\widehat{\theta}; \boldsymbol{x})} = \exp\left(-\frac{n}{2}(\theta - \widehat{\theta})^2\right).$$

It is the Gaussian fuzzy number $\mathsf{GFN}(\widehat{\theta}, n)$ with mode $\widehat{\theta}$ and precision n. Now, Y can be written as $Y = \theta + U$, with $U \sim N(0, 1)$. Consequently, the conditional possibility distribution on Y given $U = u$ is the Gaussian fuzzy number $\widetilde{\theta}_{\boldsymbol{x}} + u = \mathsf{GFN}(\widehat{\theta} + u, n)$, and our knowledge of Y is described by the random fuzzy set $U \to \mathsf{GFN}(\widehat{\theta} + U, n)$, with $U \sim N(0, 1)$. This is a Gaussian fuzzy number with fixed precision $h = n$ and normal random mode $M = \widehat{\theta} + U \sim N(\widehat{\theta}, 1)$. This important class of random fuzzy sets will be studied in the next section.

4. Gaussian random fuzzy numbers

In this section, we introduce Gaussian random fuzzy numbers (GRFN's) as a practical model for representing uncertainty on a real variable. As we will see, this model encompasses Gaussian random variables and Gaussian fuzzy numbers as special cases. A GRFN can be seen, equivalently, as a Gaussian random variable with fuzzy mean, or as a Gaussian fuzzy number with random mode. The definition and main properties will first be presented in Section 4.1. The expression of the orthogonal sum of two GRFN's will then be derived in Section 4.2. Finally, arithmetic operations on GRFN's will be addressed in Section 4.3.

4.1. Definition and main properties

Definition 1. *Let $(\Omega, \sigma_\Omega, P)$ be a probability space and let $M : \Omega \to \mathbb{R}$ be a Gaussian random variable with mean μ and variance σ^2. The random fuzzy set $\widetilde{X} : \Omega \to [0, 1]^{\mathbb{R}}$ defined as*

$$\widetilde{X}(\omega) = \mathsf{GFN}(M(\omega), h)$$

is called a Gaussian random fuzzy number (GRFN) with mean μ, variance σ^2 and precision h, which we write $\widetilde{X} \sim \widetilde{N}(\mu, \sigma^2, h)$.

In the definition of a GRFN, μ is a location parameter, while parameters h and σ^2 correspond, respectively, to possibilistic and probabilistic uncertainty. If $h = 0$, imprecision is maximal whatever the values of μ and σ^2: the GRFN $\widetilde{X}$ then induces the vacuous belief function on $\mathbb{R}$, in which case $Bel_{\widetilde{X}}(A) = 0$ for all $A \subset \mathbb{R}$, and $Pl_{\widetilde{X}}(A) = 1$ for all $A \subseteq \mathbb{R}$ such that $A \neq \emptyset$; such a GRFN will be said to be *vacuous* and will be denoted by $\widetilde{X} \sim \widetilde{N}(0, 1, 0)$. If $h = +\infty$, each fuzzy number $\mathsf{GFN}(M(\omega), h)$ is reduced to a point: the GRFN $\widetilde{X}$ is then equivalent to a Gaussian random variable with mean μ and variance σ^2, which we can write: $\widetilde{N}(\mu, \sigma^2, +\infty) = N(\mu, \sigma^2)$. Another special case of interest is that where $\sigma^2 = 0$, in which case M is a constant random variable taking value μ, and $\widetilde{X}$ is a possibilistic variable with possibility distribution $\mathsf{GFN}(\mu, h)$.

The following proposition gives the expression of the contour functions $pl_{\widetilde{X}}(x)$ associated to $\widetilde{X}$.

Proposition 6. *The contour function of GRFN $\widetilde{X} \sim \widetilde{N}(\mu, \sigma^2, h)$ is*

$$pl_{\widetilde{X}}(x) = \frac{1}{\sqrt{1 + h\sigma^2}} \exp\left(-\frac{h(x - \mu)^2}{2(1 + h\sigma^2)}\right). \tag{22}$$

Proof. See Appendix C. □

A shown by Proposition 6, the contour function $pl_{\widetilde{X}}$ is constant in two cases: if $h = 0$, $\widetilde{X}$ is vacuous, and $pl_{\widetilde{X}}(x) = 1$ for all $x \in \mathbb{R}$; if $h = +\infty$, $\widetilde{X}$ is a random variable, and $pl_{\widetilde{X}}(x) = 0$ for all $x \in \mathbb{R}$. We also note that, if $\sigma^2 = 0$, $pl_{\widetilde{X}}$ is equal to the possibility distribution $\mathsf{GFN}(\mu, h)$. When $\sigma^2 \to +\infty$ and $h > 0$, $pl_{\widetilde{X}}(x) \to 0$ for all x. The next proposition gives the expressions of the belief and plausibility of any real interval.

Proposition 7. *For any real interval $[x, y]$, the degrees of belief and plausibility of $[x, y]$ induced by the GRFN $\widetilde{X} \sim \widetilde{N}(\mu, \sigma^2, h)$ are, respectively,*

$$Bel_{\widetilde{X}}([x, y]) = \Phi\left(\frac{y - \mu}{\sigma}\right) - \Phi\left(\frac{x - \mu}{\sigma}\right) -$$
$$pl_{\widetilde{X}}(x)\left[\Phi\left(\frac{(x+y)/2 - \mu + (y-x)h\sigma^2/2}{\sigma\sqrt{h\sigma^2 + 1}}\right) - \Phi\left(\frac{x - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right)\right] -$$
$$pl_{\widetilde{X}}(y)\left[\Phi\left(\frac{y - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right) - \Phi\left(\frac{(x+y)/2 - \mu - (y-x)h\sigma^2/2}{\sigma\sqrt{h\sigma^2 + 1}}\right)\right], \quad (23)$$

and

$$Pl_{\widetilde{X}}([x, y]) = \Phi\left(\frac{y - \mu}{\sigma}\right) - \Phi\left(\frac{x - \mu}{\sigma}\right) + pl_{\widetilde{X}}(x)\Phi\left(\frac{x - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right) +$$
$$pl_{\widetilde{X}}(y)\left[1 - \Phi\left(\frac{y - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right)\right]. \quad (24)$$

Proof. See Appendix D. □

Corollary 1. *The lower and upper cdf's of the GRFN $\widetilde{X} \sim \widetilde{N}(\mu, \sigma^2, h)$ are, respectively*

$$Bel_{\widetilde{X}}((-\infty, y]) = \Phi\left(\frac{y - \mu}{\sigma}\right) - pl_{\widetilde{X}}(y)\Phi\left(\frac{y - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right) \quad (25)$$

and

$$Pl_{\widetilde{X}}((-\infty, y]) = \Phi\left(\frac{y - \mu}{\sigma}\right) + pl_{\widetilde{X}}(y)\left[1 - \Phi\left(\frac{y - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right)\right]. \quad (26)$$

Proof. Immediate from Proposition 7 by letting x tend to $-\infty$ in (23) and (24) □

We can easily check from (23) and (24) that $Bel_{\widetilde{X}}([x, y])$ and $Pl_{\widetilde{X}}([x, y])$ both tend to $\Phi\left(\frac{y - \mu}{\sigma}\right) - \Phi\left(\frac{x - \mu}{\sigma}\right)$ when $h \to \infty$, which is consistent with the fact that a GRFN with infinite precision is equivalent to a Gaussian random variable. Finally, the following proposition gives the expressions of the lower and upper expectations of a GRFN.

Proposition 8. *Let $\widetilde{X} \sim \widetilde{N}(\mu, \sigma^2, h)$ be a GRFN with $h > 0$. Its lower and upper expectations are, respectively,*

$$\mathbb{E}_*(\widetilde{X}) = \mu - \sqrt{\frac{\pi}{2h}} \quad and \quad \mathbb{E}^*(\widetilde{X}) = \mu + \sqrt{\frac{\pi}{2h}}. \quad (27)$$

22

Proof. See Appendix E. □

As expected, we can see from (27) that the lower and upper expectations boil down to the usual expectation μ when $h = +\infty$.

4.2. Orthogonal sum of Gaussian random fuzzy numbers

In this section, we derive the expression of the orthogonal sum $\widetilde{X}_1 \oplus \widetilde{X}_2$ of two GRFN's $\widetilde{X}_1$ and $\widetilde{X}_2$. We start with the following lemma.

Lemma 1. *Let $M_1 \sim N(\mu_1, \sigma_1^2)$ and $M_2 \sim N(\mu_2, \sigma_2^2)$ be two independent Gaussian random variables, and let $\widetilde{F}$ be the fuzzy subset of $\mathbb{R}^2$ with membership function*

$$\widetilde{F}(m_1, m_2) = \mathsf{hgt}\left(\mathsf{GFN}(m_1, h_1) \cdot \mathsf{GFN}(m_2, h_2)\right) = \exp\left(-\frac{h_1 h_2 (m_1 - m_2)^2}{2(h_1 + h_2)}\right).$$

The conditional probability distribution of (M_1, M_2) given $\widetilde{F}$ is two-dimensional Gaussian with mean vector $\widetilde{\boldsymbol{\mu}} = (\widetilde{\mu}_1, \widetilde{\mu}_2)^T$ and covariance matrix

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \widetilde{\sigma}_1^2 & \rho \widetilde{\sigma}_1 \widetilde{\sigma}_2 \\ \rho \widetilde{\sigma}_1 \widetilde{\sigma}_2 & \widetilde{\sigma}_2^2 \end{pmatrix},$$

with

$$\widetilde{\mu}_1 = \frac{\mu_1(1 + \overline{h}\sigma_2^2) + \mu_2 \overline{h}\sigma_1^2}{1 + \overline{h}(\sigma_1^2 + \sigma_2^2)} \tag{28a}$$

$$\widetilde{\mu}_2 = \frac{\mu_2(1 + \overline{h}\sigma_1^2) + \mu_1 \overline{h}\sigma_2^2}{1 + \overline{h}(\sigma_1^2 + \sigma_2^2)} \tag{28b}$$

$$\widetilde{\sigma}_1^2 = \frac{\sigma_1^2(1 + \overline{h}\sigma_2^2)}{1 + \overline{h}(\sigma_1^2 + \sigma_2^2)} \tag{28c}$$

$$\widetilde{\sigma}_2^2 = \frac{\sigma_2^2(1 + \overline{h}\sigma_1^2)}{1 + \overline{h}(\sigma_1^2 + \sigma_2^2)} \tag{28d}$$

$$\rho = \frac{\overline{h}\sigma_1\sigma_2}{\sqrt{(1 + \overline{h}\sigma_1^2)(1 + \overline{h}\sigma_2^2)}}, \tag{28e}$$

where

$$\overline{h} = \frac{h_1 h_2}{h_1 + h_2}. \tag{28f}$$

Furthermore, the degree of conflict between two independent GRFN's $\widetilde{X}_1 \sim \widetilde{N}(\mu_1, \sigma_1^2, h_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, \sigma_2^2, h_2)$ is

$$\kappa = 1 - \iint f(m_1, m_2)\widetilde{F}(m_1, m_2) dm_1 dm_2 =$$

$$1 - \frac{\widetilde{\sigma}_1 \widetilde{\sigma}_2}{\sigma_1 \sigma_2}\sqrt{1 - \rho^2} \exp\left\{ -\frac{1}{2}\left[\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right] + \frac{1}{2(1 - \rho^2)}\left[\frac{\widetilde{\mu}_1^2}{\widetilde{\sigma}_1^2} + \frac{\widetilde{\mu}_2^2}{\widetilde{\sigma}_2^2} - 2\rho\frac{\widetilde{\mu}_1 \widetilde{\mu}_2}{\widetilde{\sigma}_1 \widetilde{\sigma}_2}\right]\right\},$$

where $f(m_1, m_2)$ is the pdf of random vector (M_1, M_2).

Proof. See Appendix F. □

Proposition 9. *Let $\widetilde{X}_1 \sim \widetilde{N}(\mu_1, \sigma_1^2, h_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, \sigma_2^2, h_2)$ be two independent GRFN's, and assume that $h_1 > 0$ or $h_2 > 0$. We have*

$$\widetilde{X}_1 \oplus \widetilde{X}_2 \sim \widetilde{N}(\widetilde{\mu}_{12}, \widetilde{\sigma}_{12}^2, h_{12}),$$

with

$$h_{12} = h_1 + h_2, \tag{29}$$

$$\widetilde{\mu}_{12} = \frac{h_1 \widetilde{\mu}_1 + h_2 \widetilde{\mu}_2}{h_1 + h_2}, \tag{30}$$

and

$$\widetilde{\sigma}_{12}^2 = \frac{h_1^2 \widetilde{\sigma}_1^2 + h_2^2 \widetilde{\sigma}_2^2 + 2\rho h_1 h_2 \widetilde{\sigma}_1 \widetilde{\sigma}_2}{(h_1 + h_2)^2}, \tag{31}$$

where $\widetilde{\mu}_1$, $\widetilde{\mu}_2$, $\widetilde{\sigma}_1$, $\widetilde{\sigma}_2$ and ρ are given by (28) in Lemma 1.

Proof. Let M_1 and M_2 be the Gaussian random variables from $(\Omega_1, \sigma_1, P_1)$ and $(\Omega_2, \sigma_2, P_2)$ to $(\mathbb{R}, \beta_{\mathbb{R}})$ corresponding, respectively, to GRFN's $\widetilde{X}_1 \sim \widetilde{N}(\mu_1, \sigma_1^2, h_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, \sigma_2^2, h_2)$. The orthogonal sum of $\widetilde{X}_1$ and $\widetilde{X}_2$ is the random fuzzy set $(\Omega_1 \times \Omega_2, \sigma_1 \otimes \sigma_2, \widetilde{P}_{12}, \mathbb{R}, \beta_{\mathbb{R}}, \widetilde{X}_\odot)$, where $\widetilde{X}_\odot$ is the mapping

$$\widetilde{X}_\odot : (\omega_1, \omega_2) \to \mathsf{GFN}(M_{12}(\omega_1, \omega_2), h_1 + h_2),$$

with

$$M_{12}(\omega_1, \omega_2) = \frac{h_1 M_1(\omega_1) + h_2 M_2(\omega_2)}{h_1 + h_2},$$

and $\widetilde{P}_{12}$ is the probability measure on $\Omega_1 \times \Omega_2$ obtained by conditioning $P_1 \times P_2$ on the fuzzy set $\widetilde{\Theta}^*(\omega_1, \omega_2) = \mathsf{hgt}\left(\mathsf{GFN}(M_1(\omega_1), h_1), \mathsf{GFN}(M_2(\omega_2), h_2)\right)$. From Lemma 1, the pushforward measure of $\widetilde{P}_{12}$ by the random vector (M_1, M_2) is the two-dimensional Gaussian distribution with parameters $(\widetilde{\mu}_1, \widetilde{\mu}_2, \widetilde{\sigma}_1, \widetilde{\sigma}_2, \rho)$. Consequently, M_{12} is a Gaussian random variable with mean

$$\mathbb{E}(M_{12}) = \frac{h_1 \mathbb{E}(M_1) + h_2 \mathbb{E}(M_2)}{h_1 + h_2} = \frac{h_1 \widetilde{\mu}_1 + h_2 \widetilde{\mu}_2}{h_1 + h_2},$$

and variance

$$\begin{aligned} \mathrm{Var}(M_{12}) &= \frac{h_1^2 \mathrm{Var}(M_1) + h_2^2 \mathrm{Var}(M_2) + 2h_1 h_2 \mathrm{Cov}(M_1, M_2)}{(h_1 + h_2)^2} \\ &= \frac{h_1^2 \widetilde{\sigma}_1^2 + h_2^2 \widetilde{\sigma}_2^2 + 2\rho h_1 h_2 \widetilde{\sigma}_1 \widetilde{\sigma}_2}{(h_1 + h_2)^2}. \end{aligned}$$

□

24

Let us now consider some special cases in which one of two GRFN's is a Gaussian random variable. The next proposition states that the orthogonal sum of a Gaussian random variable and an arbitrary GRFN with finite precision is a Gaussian random variable.

Proposition 10. *Let $X_1 \sim N(\mu_1, \sigma_1^2)$ be a Gaussian random variable and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, \sigma_2^2, h_2)$ a GRFN with finite precision $h_2 < +\infty$. Their orthogonal sum is a Gaussian random variable $X_1 \oplus \widetilde{X}_2 \sim N(\widetilde{\mu}_{12}, \widetilde{\sigma}_{12}^2)$ with*

$$\widetilde{\mu}_{12} = \frac{\mu_1(1 + h_2\sigma_2^2) + \mu_2 h_2\sigma_1^2}{1 + h_2(\sigma_1^2 + \sigma_2^2)}, \tag{32}$$

$$\widetilde{\sigma}_{12}^2 = \frac{\sigma_1^2(1 + h_2\sigma_2^2)}{1 + h_2(\sigma_1^2 + \sigma_2^2)}, \tag{33}$$

and the probability density of $X_1 \oplus \widetilde{X}_2$ is proportional to the product of the pdf of X_1 and the contour function of $\widetilde{X}_2$.

Proof. See Appendix G. $\qquad \square$

The following corollary addresses the special case where $\widetilde{X}_2$ is a possibilistic GRFN.

Corollary 2. *Let $X_1 \sim N(\mu_1, \sigma_1^2)$ be a Gaussian random variable and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, 0, h_2)$ a possibilistic GRFN. Their orthogonal sum $X_1 \oplus \widetilde{X}_2$ is a Gaussian random variable and its distribution is the conditional distribution of X_1 given the fuzzy event $\mathsf{GFN}(\mu_2, h_2)$.*

Proof. From Proposition 10, $X_1 \oplus \widetilde{X}_2 \sim \widetilde{N}(\widetilde{\mu}_{12}, \widetilde{\sigma}_{12}^2)$ with

$$\widetilde{\mu}_{12} = \frac{\mu_1 + \mu_2 h_2\sigma_1^2}{1 + h_2\sigma_1^2} \quad \text{and} \quad \widetilde{\sigma}_{12}^2 = \frac{\sigma_1^2}{1 + h_2\sigma_1^2}.$$

Now, we know from Proposition 10 that the density of $X_1 \oplus \widetilde{X}_2$ is proportional to the product of the density of X_1 and the contour function of $\widetilde{X}_2$, which is $\varphi(x; \mu_2, h_2)$. Consequently, we have

$$f_{X_1 \oplus \widetilde{X}_2}(x) = \frac{\frac{1}{\sigma_1^2}\exp\left(-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right)\exp\left(-\frac{h_2(x-\mu_2)^2}{2(1+h_2\sigma_2^2)}\right)}{\int \frac{1}{\sigma_1^2}\exp\left(-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right)\exp\left(-\frac{h_2(x-\mu_2)^2}{2(1+h_2\sigma_2^2)}\right)dx},$$

which is the conditional density $f_{X_1}(x|\mathsf{GFN}(\mu_2, h_2))$. $\qquad \square$

Finally, another special case of interest is when both GRFN's are Gaussian random variables. This case is addressed by the following corollary.

Corollary 3. *Let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ be two Gaussian random variables. We have $X_1 \oplus X_2 \sim N(\widetilde{\mu}_{12}, \sigma_{12}^2)$ with*

$$\widetilde{\mu}_{12} = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \quad \text{and} \quad \widetilde{\sigma}_{12}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

Proof. Immediate from Proposition 10 by letting h_2 tend to $+\infty$ in (32) and (33). $\qquad \square$

4.3. Arithmetic operations on GRFN's

Arithmetic operations can be extended to fuzzy numbers using Zadeh's extension principles [14, 12]. More precisely, let $\widetilde{A}$ and $\widetilde{B}$ be two fuzzy numbers, and let $*$ be a binary operation on reals. Then the fuzzy number $\widetilde{C} = \widetilde{A} * \widetilde{B}$ is defined as

$$\widetilde{C}(c) = \sup_{c=a*b} \min(\widetilde{A}(a), \widetilde{B}(b)).$$

The membership function $\widetilde{C}$ is equal to the possibility distribution on $c = a * b$, if a and b are constrained, respectively, by possibility distributions $\widetilde{A}$ and $\widetilde{B}$. Unary or n-ary operations can be extended from real to fuzzy numbers in the same way. For a certain class of fuzzy number called LR-fuzzy numbers [14, page 54], closed-form expressions exist for the addition, subtraction and scalar multiplication of fuzzy numbers. In particular, Gaussian fuzzy numbers with positive precision are LR fuzzy numbers and they verify the following equalities [25]:

$$\mathsf{GFN}(m_1, h_1) + \mathsf{GFN}(m_2, h_2) = \mathsf{GFN}(m_1 + m_2, (h_1^{-1/2} + h_2^{-1/2})^{-2})$$
$$\mathsf{GFN}(m_1, h_1) - \mathsf{GFN}(m_2, h_2) = \mathsf{GFN}(m_1 - m_2, (h_1^{-1/2} + h_2^{-1/2})^{-2})$$
$$\lambda \cdot \mathsf{GFN}(m, h) = \mathsf{GFN}(\lambda m, h/\lambda^2), \quad \forall \lambda \in \mathbb{R}.$$

As addition of fuzzy numbers is associative, we can express the linear combination of n GFN's as

$$\sum_{i=1}^{n} \lambda_i \cdot \mathsf{GFN}(m_i, h_i) = \mathsf{GFN}\left(\sum_{i=1}^{n} \lambda_i m_i, \left(\sum_{i=1}^{n} |\lambda_i| h_i^{-1/2}\right)^{-2}\right). \tag{34}$$

Now, let us consider n independent GRFN's $\widetilde{X}_i$ from probability spaces $(\Omega_i, \sigma_i, P_i)$ to $[0,1]^{\mathbb{R}}$ defined by

$$\widetilde{X}_i(\omega) = \mathsf{GFN}(M_i(\omega), h_i)$$

for all $\omega \in \Omega_i$, where M_i is a Gaussian random variable with mean μ_i and standard deviation σ_i, and $h_i > 0$. Let

$$\widetilde{X} = \sum_{i=1}^{n} \lambda_i \widetilde{X}_i$$

be the random fuzzy set from $(\Omega_1 \times \ldots \times \Omega_n, \sigma_1 \otimes \ldots \otimes \sigma_n, P_1 \times \ldots \times P_n)$ to $[0,1]^{\mathbb{R}}$ defined by

$$\widetilde{X}(\omega_1, \ldots, \omega_n) = \sum_{i=1}^{n} \lambda_i \cdot \mathsf{GFN}(M_i(\omega_i), h_i).$$

If each GRFN $\widetilde{X}_i$ represents our knowledge about the value of some quantity X_i, $\widetilde{X}$ represents our knowledge about $X = \sum_{i=1}^{n} \lambda_i X_i$. From (34), $\widetilde{X} \sim \widetilde{N}(\mu, \sigma, h)$ with

$$\mu = \sum_{i=1}^{n} \lambda_i \mu_i, \quad \sigma^2 = \sum_{i=1}^{n} \lambda_i^2 \sigma_i^2, \quad \text{and} \quad h = \left(\sum_{i=1}^{n} |\lambda_i| h_i^{-1/2}\right)^{-2}.$$

5. Gaussian random fuzzy vectors

In this section, we introduce Gaussian random fuzzy vectors (GRFV's), an extension of the model presented in Section 4 allowing us to describe knowledge about multidimensional quantities. The main definitions and properties are first introduced in Section 5.1. The expression of the orthogonal sum of two GRFV's is then given in Section 5.2, after which the marginalization and vacuous extension of GRFV's are described in Section 5.3. Finally, our model is compared to Dempster's normal belief function model in Section 5.4.

5.1. Definition and main properties

We consider a p-dimensional variable $\boldsymbol{\theta}$ taking values in $\mathbb{R}^p$. Knowledge about $\boldsymbol{\theta}$ may be encoded as a p-dimensional Gaussian fuzzy vector, defined as follows.

Definition 2. *We define the p-dimensional Gaussian fuzzy vector (GFV) with center $\boldsymbol{m} \in \mathbb{R}^p$ and $p \times p$ symmetric and positive semidefinite precision matrix $\boldsymbol{H}$ as the normalized fuzzy subset of $\mathbb{R}^p$ with membership function*

$$\varphi(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{H}) = \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{m})\right),$$

denoted as GFV($\boldsymbol{m}, \boldsymbol{H}$).

As shown in [27], the normalized product of two GFV's is still a GFV. The following proposition generalizes Proposition 3.

Proposition 11. *Let GFV($\boldsymbol{m}_1, \boldsymbol{H}_1$) and GFV($\boldsymbol{m}_2, \boldsymbol{H}_2$) be two p-dimensional GFV's with positive definite precision matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$. We have*

$$\varphi(\boldsymbol{x}; \boldsymbol{m}_1, \boldsymbol{H}_1) \cdot \varphi(\boldsymbol{x}; \boldsymbol{m}_2, \boldsymbol{H}_2) = \varphi(\boldsymbol{x}; \boldsymbol{m}_{12}, \boldsymbol{H}_{12}) \times$$
$$\exp\left(-\frac{1}{2}(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T (\boldsymbol{H}_1^{-1} + \boldsymbol{H}_2^{-1})^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2)\right),$$

with
$$\boldsymbol{m}_{12} = (\boldsymbol{H}_1 + \boldsymbol{H}_2)^{-1}(\boldsymbol{H}_1 \boldsymbol{m}_1 + \boldsymbol{H}_2 \boldsymbol{m}_2) \quad and \quad \boldsymbol{H}_{12} = \boldsymbol{H}_1 + \boldsymbol{H}_2.$$

Consequently, the following equation holds:

$$GFV(\boldsymbol{m}_1, \boldsymbol{H}_1) \odot GFV(\boldsymbol{m}_2, \boldsymbol{H}_2) = GFV(\boldsymbol{m}_{12}, \boldsymbol{H}_{12}),$$

and the height of the product intersection between GFV($\boldsymbol{m}_1, \boldsymbol{H}_1$) and GFV($\boldsymbol{m}_1, \boldsymbol{H}_2$) is

$$hgt\left(GFV(\boldsymbol{m}_1, \boldsymbol{H}_1), GFV(\boldsymbol{m}_1, \boldsymbol{H}_2)\right) = \max_{\boldsymbol{x}} \varphi(\boldsymbol{x}; \boldsymbol{m}_1, \boldsymbol{H}_1)\varphi(\boldsymbol{x}; \boldsymbol{m}_2, \boldsymbol{H}_2) \tag{35a}$$

$$= \exp\left(-\frac{1}{2}(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T (\boldsymbol{H}_1^{-1} + \boldsymbol{H}_2^{-1})^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2)\right). \tag{35b}$$

Equipped with the notion of GFV, we can now introduce a model of random fuzzy set that can be seen as a GFV whose mode is a multidimensional Gaussian random variable. This model is defined formally as follows.

Definition 3. *Let $(\Omega, \sigma_\Omega, P)$ be a probability space, $\boldsymbol{M} : \Omega \to \mathbb{R}^p$ a p-dimensional Gaussian random vector with mean $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Sigma}$, and $\boldsymbol{H}$ a $p \times p$ symmetric and positive semidefinite real matrix. The random fuzzy set $\widetilde{X} : \Omega \to [0,1]^{\mathbb{R}^p}$ defined as*

$$\widetilde{X}(\omega) = \mathsf{GFV}(\boldsymbol{M}(\omega), \boldsymbol{H})$$

is called a Gaussian random fuzzy vector (GRFV), which we denote as $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$.

The following proposition generalizes Proposition 6.

Proposition 12. *The contour function of GRFV $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$ with positive definite precision matrix $\boldsymbol{H}$ is*

$$pl_{\widetilde{X}}(\boldsymbol{x}) = \frac{1}{|\boldsymbol{I}_p + \boldsymbol{\Sigma}\boldsymbol{H}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T(\boldsymbol{H}^{-1} + \boldsymbol{\Sigma})^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right),$$

where $\boldsymbol{I}_p$ is the p-dimensional identity matrix.

Proof. See Appendix H. $\qquad\square$

5.2. Orthogonal sum of Gaussian random fuzzy vectors

The practical interest of GRFV's arises from the fact that they can be easily combined by the generalized product-intersection rule. The following lemma and proposition, which generalize, respectively, Lemma 1 and Proposition 9, give us the expression of the orthogonal sum of two GRFV's.

Lemma 2. *Let $\boldsymbol{M}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\boldsymbol{M}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ be two independent Gaussian p-dimensional random vectors and let $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ be two symmetric and positive definite $p \times p$ matrices. Let $\widetilde{F}$ be the fuzzy subset of $\mathbb{R}^{2p}$ with membership function*

$$\widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) = \mathsf{hgt}\left(\mathsf{GFV}(\boldsymbol{m}_1, \boldsymbol{H}_1) \cdot \mathsf{GFV}(\boldsymbol{m}_2, \boldsymbol{H}_2)\right),$$

and let $\boldsymbol{M}$ be the 2p-dimensional vector $(\boldsymbol{M}_1, \boldsymbol{M}_2)$. The conditional probability distribution of $\boldsymbol{M}$ given $\widetilde{F}$ is 2p-dimensional Gaussian with mean vector $\widetilde{\boldsymbol{\mu}}$ and covariance matrix $\widetilde{\boldsymbol{\Sigma}}$ defined as follows:

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Sigma}_1^{-1} + \overline{\boldsymbol{H}} & -\overline{\boldsymbol{H}} \\ -\overline{\boldsymbol{H}} & \boldsymbol{\Sigma}_2^{-1} + \overline{\boldsymbol{H}} \end{pmatrix}^{-1},$$

$$\widetilde{\boldsymbol{\mu}} = \begin{pmatrix} \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{I}_p & -\boldsymbol{I}_p \\ -\boldsymbol{I}_p & \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1} + \boldsymbol{I}_p \end{pmatrix}^{-1} \begin{pmatrix} \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad (36)$$

with

$$\overline{\boldsymbol{H}} = (\boldsymbol{H}_1^{-1} + \boldsymbol{H}_2^{-1})^{-1}.$$

Furthermore, the degree of conflict between two GRFV's $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{H}_1)$ *and* $\widetilde{X}_2 \sim \widetilde{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{H}_2)$ *is*

$$\kappa = 1 - \int_{\mathbb{R}^{2p}} f(\boldsymbol{m}_1, \boldsymbol{m}_2) \widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) d\boldsymbol{m}_1 d\boldsymbol{m}_2 =$$

$$1 - \sqrt{\frac{|\widetilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}} \exp\left\{ -\frac{1}{2} \left[\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\mu}_2 - \widetilde{\boldsymbol{\mu}}^T \widetilde{\boldsymbol{\Sigma}}^{-1} \widetilde{\boldsymbol{\mu}} \right] \right\}.$$

Proof. See Appendix I □

Proposition 13. *Let* $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{H}_1)$ *and* $\widetilde{X}_2 \sim \widetilde{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{H}_2)$ *be two independent GRFV's. We have*

$$\widetilde{X}_1 \oplus \widetilde{X}_2 \sim \widetilde{N}(\widetilde{\boldsymbol{\mu}}_{12}, \widetilde{\boldsymbol{\Sigma}}_{12}, \boldsymbol{H}_{12})$$

with

$$\boldsymbol{H}_{12} = \boldsymbol{H}_1 + \boldsymbol{H}_2,$$
$$\widetilde{\boldsymbol{\mu}}_{12} = \boldsymbol{A}\widetilde{\boldsymbol{\mu}},$$

and

$$\widetilde{\boldsymbol{\Sigma}}_{12} = \boldsymbol{A}\widetilde{\boldsymbol{\Sigma}}\boldsymbol{A}^T,$$

where $\boldsymbol{A}$ *is the constant* $p \times 2p$ *matrix defined as*

$$\boldsymbol{A} = \boldsymbol{H}_{12}^{-1} \begin{pmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 \end{pmatrix}.$$

Proof. Let $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ be the Gaussian random vector from $(\Omega_1, \sigma_1, P_1)$ and $(\Omega_2, \sigma_2, P_2)$ to $(\mathbb{R}^p, \beta_{\mathbb{R}^p})$ corresponding, respectively, to GRFV's $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{H}_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{H}_2)$. The orthogonal sum of $\widetilde{X}_1$ and $\widetilde{X}_2$ is defined by the mapping

$$\widetilde{X}_{\odot} : (\omega_1, \omega_2) \to \mathsf{GFV}(\boldsymbol{M}_{12}(\omega_1, \omega_2), \boldsymbol{H}_1 + \boldsymbol{H}_2)$$

with

$$\boldsymbol{M}_{12} = (\boldsymbol{H}_1 + \boldsymbol{H}_2)^{-1}(\boldsymbol{H}_1\boldsymbol{M}_1 + \boldsymbol{H}_2\boldsymbol{M}_2) = \boldsymbol{A} \begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \end{pmatrix},$$

where $\boldsymbol{A}$ is the $p \times 2p$ matrix

$$\boldsymbol{A} = (\boldsymbol{H}_1 + \boldsymbol{H}_2)^{-1} \begin{pmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 \end{pmatrix},$$

and the probability measure $\widetilde{P}_{12}$ on $\Omega_1 \times \Omega_2$ obtained by conditioning $P_1 \times P_2$ on the fuzzy set $\widetilde{\Theta}^*(\omega_1, \omega_2) = \mathsf{hgt}\left(\mathsf{GFV}(\boldsymbol{M}_1(\omega_1), \boldsymbol{H}_1), \mathsf{GFV}(\boldsymbol{M}_2(\omega_2), \boldsymbol{H}_2)\right)$. From Lemma 2, the pushforward measure of $\widetilde{P}_{12}$ by the random vector $(\boldsymbol{M}_1, \boldsymbol{M}_2)$ is the p-dimensional Gaussian distribution with parameters $(\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}})$. Consequently, $\boldsymbol{M}_{12}$ is a Gaussian random vector with mean

$$\mathbb{E}(\boldsymbol{M}_{12}) = \boldsymbol{A}\widetilde{\boldsymbol{\mu}}$$

and variance

$$\mathrm{Var}(\boldsymbol{M}_{12}) = \boldsymbol{A}\widetilde{\boldsymbol{\Sigma}}\boldsymbol{A}^T.$$

□

5.3. Marginalization and vacuous extension

In this section, we consider the marginalization and vacuous extension (defined in Section 3.3) of a GRFV. We assume that variable $\boldsymbol{\theta}$ taking values in $\mathbb{R}^p$ is decomposed as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ with $\boldsymbol{\theta}_1 \in \Theta_1 = \mathbb{R}^{p-k}$ and $\boldsymbol{\theta}_2 \in \Theta_2 = \mathbb{R}^k$ for $0 < k < p$.

Marginalization. We start with the following lemma.

Lemma 3. *Let $\widetilde{F} = \mathsf{GFV}(\boldsymbol{m}, \boldsymbol{H})$ be a p-dimensional Gaussian fuzzy vector with mode $\boldsymbol{m} = (\boldsymbol{m}_1, \boldsymbol{m}_2)$, where $\boldsymbol{m}_1 \in \Theta_1 = \mathbb{R}^{p-k}$ and $\boldsymbol{m}_2 \in \Theta_2 = \mathbb{R}^k$ for $0 < k < p$, and precision matrix $\boldsymbol{H}$ with block decomposition*

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{pmatrix}.$$

Assume that $\boldsymbol{H}_{22}$ is nonsingular. The projection of $\widetilde{F}$ on Θ_1, denoted as $\widetilde{F} \downarrow \Theta_1$ is the Gaussian fuzzy vector $\mathsf{GFV}(\boldsymbol{m}_1, \boldsymbol{H}'_{11})$ with

$$\boldsymbol{H}'_{11} = \boldsymbol{H}_{11} - \boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}.$$

Proof. See Appendix J $\hfill\square$

Let us now consider a p-dimensional GRFV $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$ representing partial knowledge about $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. The marginal RFS for $\boldsymbol{\theta}_1$ is given by the following proposition, which follows directly from Lemma 3.

Proposition 14. *Let $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$ by a p-dimensional GRFV taking values in 2^Θ, with $\Theta = \Theta_1 \times \Theta_2$, where $\Theta_1 = \mathbb{R}^{p-k}$ and $\Theta_2 = \mathbb{R}^k$ for $0 < k < p$. Let $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ with $\boldsymbol{\mu}_1 \in \Theta_1$ and $\boldsymbol{\mu}_2 \in \Theta_2$, and consider the block decompositions*

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \quad and \quad \boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{pmatrix}.$$

Assume that $\boldsymbol{H}_{22}$ is nonsingular. The marginal of $\widetilde{X}$ on Θ_1 is the GRFV $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \boldsymbol{H}'_{11})$ with

$$\boldsymbol{H}'_{11} = \boldsymbol{H}_{11} - \boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}.$$

Vacuous extension. We now consider a Gaussian fuzzy vector $\mathsf{GFV}(\boldsymbol{m}_1, \boldsymbol{H}_{11})$ in $\Theta_1 = \mathbb{R}^{p-k}$ for $0 < k < p$. Its cylindrical extension in $\Theta = \Theta_1 \times \Theta_2$, with $\Theta_2 = \mathbb{R}^k$ has the following membership function

$$\varphi(\boldsymbol{x}) = \exp\left(-\frac{1}{2}(\boldsymbol{x}_1 - \boldsymbol{m}_1)^T \boldsymbol{H}_{11}(\boldsymbol{x}_1 - \boldsymbol{m}_1)\right),$$

which can be written as

$$\varphi(\boldsymbol{x}) = \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{m})\right),$$

where $\boldsymbol{m}$ is the p-dimensional vector

$$\boldsymbol{m} = \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{0} \end{pmatrix}$$

and $\boldsymbol{H}$ is the $p \times p$ matrix

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{37}$$

Given a GRFV $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \boldsymbol{H}_{11})$ taking values in 2^{Θ_1}, it follows immediately that its vacuous extension in $\Theta = \Theta_1 \times \Theta_2$ is the GRFV

$$\widetilde{X}_{1\uparrow(1,2)} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$$

with

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_k \end{pmatrix},$$

where $\boldsymbol{I}_k$ is the $k \times k$ identity matrix, and $\boldsymbol{H}$ given by (37).

Noninteractivity. In Section 3.3, we defined the notion of noninteractive random fuzzy vector. The following proposition gives a necessary and sufficient condition for a GRFV to be noninteractive.

Proposition 15. *A p-dimensional GRFV $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{H})$ is noninteractive iff matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{H}$ are diagonal.*

Proof. Let $\widetilde{X}_1, \ldots, \widetilde{X}_p$ be the marginals of $\widetilde{X}$ on each of the p coordinates. Let $\sigma_1^2, \ldots, \sigma_p^2$ and $h_1, \ldots, h_p$ be the diagonal elements of, respectively, $\boldsymbol{\Sigma}$ and $\boldsymbol{H}$. Let Ω be the set of departure of $\widetilde{X}$. Let $\widetilde{X}_{i\uparrow(1:p)}$ denote the vacuous extension of $\widetilde{X}_i$ in $\mathbb{R}^p$, defined by

$$\widetilde{X}_{i\uparrow(1:p)}(\omega)(\boldsymbol{x}) = \exp\left(-\frac{h}{2}(x_i - M_i(\omega))^2\right)$$

with $M_i \sim N(\mu_i, \sigma_i^2)$. The orthogonal sum

$$\widetilde{X}' = \widetilde{X}_{1\uparrow(1:p)} \oplus \ldots \oplus \widetilde{X}_{p\uparrow(1:p)}$$

is given by

$$\widetilde{X}'(\omega)(\boldsymbol{x}) = \prod_{i=1}^{p} \exp\left(-\frac{h}{2}(x_i - M_i(\omega))^2\right) = \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{M}'(\omega))^T \boldsymbol{H}'(\boldsymbol{x} - \boldsymbol{M}'(\omega))\right),$$

where $\boldsymbol{H}'$ is the diagonal matrix with diagonal elements $h_1, \ldots, h_p$, and $\boldsymbol{M}'$ is a random vector with mean $\boldsymbol{\mu}$ and diagonal covariance matrix $\boldsymbol{\Sigma}'$ with diagonal elements $\sigma_1^2, \ldots, \sigma_p^2$. We have $\widetilde{X} = \widetilde{X}'$ iff $\boldsymbol{H} = \boldsymbol{H}'$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$, i.e., if both $\boldsymbol{H}$ and $\boldsymbol{\Sigma}$ are diagonal. $\qquad\square$

5.4. Comparison with Dempster's normal belief functions

In [5], Dempster introduced another class of continuous belief functions in $\mathbb{R}^p$, called *normal belief functions*[2]. It is interesting to compare Dempster's model with ours, as both models generalize the multivariate Gaussian distribution. A normal belief function Bel on $\mathbb{R}^p$ as defined in [5] is specified by the following components:

- An n-dimensional subspace $\mathcal{S}$ of $\mathbb{R}^p$;

- A q-dimensional partition Π of $\mathcal{S}$ into parallel $n - q$ dimensional subspaces; (If $q = 0$, $\Pi = \{\mathcal{S}\}$);

- A full-rank q-dimensional Gaussian distribution $N(\mu, \Sigma)$ on Π if $q > 0$, or the discrete probability measure with mass function $m(\mathcal{S}) = 1$ if $q = 0$.

Belief function Bel is then induced by a random set from Π, equipped with the normal distribution $N(\mu, \Sigma)$ if $q > 0$ or probability mass function m if $q = 0$, to the corresponding family of parallel $n - q$ dimensional subspaces of $\mathcal{S}$. The following special cases are of interest:

1. If $p = n = q$, Bel is a Gaussian probability distribution on $\mathbb{R}^p$;
2. If $p > n = q$, Bel is a Gaussian probability distribution limited to an n-dimensional subspace of $\mathbb{R}^p$;
3. If $p = n$ and $q = 0$, Bel is vacuous;
4. If $q = 0$ while $p > n > 0$, Bel is logical with $\mathcal{S}$ as its only focal set; it is then equivalent to specifying $p - n$ linear equations;
5. If $n = q = 0$, the true point in $\mathbb{R}^p$ is known with certainty.

Like GRFV's, Dempster's normal belief functions thus include the vacuous belief function, Gaussian probability distributions, as well as vacuous extensions of marginal Gaussian distributions. However, the two models are clearly distinct. Dempster's model is based on the combination of Gaussian probability distributions and linear equations, and is specially useful in relation with linear statistical models such as the Kalman filter [5] or linear regression [24]. In contrast, in the GRFV model, focal sets are fuzzy subsets of $\mathbb{R}^n$ ($n \leq p$) with Gaussian membership functions, or cylindrical extensions of such fuzzy subsets. This model allows us to represent not only probabilistic and logical evidence, but also fuzzy information. In particular, it includes Gaussian probability distribution and Gaussian possibility distributions as special cases. We could attempt to design an even more general model that would contain both Dempster's normal belief functions and belief functions induced by GRFV's as special cases. Such a model would allow us to reason with Gaussian probability and possibility distributions as well as with linear equations. The rigorous development of such a model is left for further research.

---

[2]Ref. [5] was actually available as a working paper from the Statistical Department of Harvard University since 1990, but it only appeared as a book chapter in 2001.

6. Conclusions

In this paper, continuing a study started in [9] with the finite case, we have introduced a theory of epistemic random fuzzy sets in a general setting. An epistemic random fuzzy set represents a piece of evidence, which may be crisp or fuzzy. This framework generalizes both epistemic random sets as considered in the Dempster-Shafer theory of belief functions, and possibility distributions considered in possibility theory. Independent epistemic random fuzzy sets are combined by the generalized product-intersection rule, which extends both Dempster's rule for combining belief functions and the product intersection rule for combining possibility distributions.

In addition, we have also introduced Gaussian random fuzzy numbers (GRFN's) and their multidimensional extensions, Gaussian random fuzzy vectors (GRFV's) as practical models of random fuzzy subsets of, respectively, $\mathbb{R}$ and $\mathbb{R}^p$ with $p \geq 2$. A GRFN is described by three parameters: its mode m, its variance σ^2 and its precision h. In this setting, a Gaussian random variable can be seen as an infinitely precise GRFN ($h = +\infty$), while a Gaussian possibility distribution is a constant GRFN ($\sigma^2 = 0$). A maximally imprecise GRFN such that $h = 0$ is said to be vacuous: it represents complete ignorance. In GRFV's, the mode becomes a p-dimensional vector, while the variance and precision become positive semi-definite $p \times p$ square matrices. The practical convenience of GRFN's and GRFV's arises from the fact that they can easily be combined by the generalized product-intersection rule. Also, formulas for the projection and marginal extension fo GRFV's have been derived.

This work opens up several perspectives. Using random fuzzy sets and, in particular, GRFN's to represent expert knowledge about numerical quantities will require the development of adequate elicitation procedures. We also consider using this framework in machine learning, to quantify prediction uncertainty in regression problems. Finally, the extension of the model introduced in this paper to take into account linear equations, as well as the development of computational procedures for reasoning with GRFV's over many variables are promising avenues for further research.

References

[1] P. A. Bromiley. Products and convolutions of Gaussian probability density functions. Technical Report 2003-003, TINA, 2014. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.3007&rep=rep1&type=pdf.

[2] I. Couso and L. Sánchez. Upper and lower probabilities induced by a fuzzy random variable. *Fuzzy Sets and Systems*, 165(1):1–23, 2011.

[3] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[4] A. P. Dempster. Upper and lower probabilities generated by a random closed interval. *Annals of Mathematical Statistics*, 39(3):957–966, 1968.

[5] A. P. Dempster. Normal belief functions and the Kalman filter. In A. K. M. E. Saleh, editor, *Data Analysis from Statistical Foundations*, pages 65–68. Nova Science Publishers, 2001.

[6] T. Denœux. Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179:1362–1376, 2009.

[7] T. Denœux. Likelihood-based belief function: justification and some extensions to low-quality data. *International Journal of Approximate Reasoning*, 55(7):1535–1547, 2014.

[8] T. Denœux. Rejoinder on "likelihood-based belief function: Justification and some extensions to low-quality data". *International Journal of Approximate Reasoning*, 55(7):1614–1617, 2014.

[9] T. Denœux. Belief functions induced by random fuzzy sets: A general framework for representing uncertain and fuzzy evidence. *Fuzzy Sets and Systems*, 424:63–91, 2021.

[10] T. Denœux, D. Dubois, and H. Prade. Representations of uncertainty in artificial intelligence: Beyond probability and possibility. In P. Marquis, O. Papini, and H. Prade, editors, *A Guided Tour of Artificial Intelligence Research*, volume 1, chapter 4, pages 119–150. Springer Verlag, 2020.

[11] T. Denœux, D. Dubois, and H. Prade. Representations of uncertainty in artificial intelligence: Probability and possibility. In P. Marquis, O. Papini, and H. Prade, editors, *A Guided Tour of Artificial Intelligence Research*, volume 1, chapter 3, pages 69–117. Springer Verlag, 2020.

[12] D. Dubois, E. Kerre, R. Mesiar, and H. Prade. Fuzzy interval analysis. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy sets*, pages 483–581. Kluwer Academic Publishers, Boston, 2000.

[13] D. Dubois, H. T. Nguyen, and H. Prade. Possibility theory, probability and fuzzy sets: Misunderstandings, bridges and gaps. In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy sets*, pages 343–438. Kluwer Academic Publishers, Boston, 2000.

[14] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York, 1980.

[15] D. Dubois and H. Prade. Possibility theory: qualitative and quantitative aspects. In D. M. Gabbay and P. Smets, editors, *Handbook of Defeasible reasoning and uncertainty management systems*, volume 1, pages 169–226. Kluwer Academic Publishers, Dordrecht, 1998.

[16] D. Dubois, H. Prade, and R. Yager. Merging fuzzy information. In J. C. Bezdek, D. Dubois, and H. Prade, editors, *Fuzzy sets in approximate reasoning and information systems*, pages 335–401. Kluwer Academic Publishers, Boston, 1999.

[17] M. A. Gil, M. López-Díaz, and D. A. Ralescu. Overview on the development of fuzzy random variables. *Fuzzy Sets and Systems*, 157(19):2546–2557, 2006.

[18] P. R. Halmos. *Measure theory*. Springer Science, New York, 1950.

[19] O. Kanjanatarakul, S. Sriboonchitta, and T. Denœux. Forecasting using belief functions: an application to marketing econometrics. *International Journal of Approximate Reasoning*, 55(5):1113–1128, 2014.

[20] O. Kanjanatarakul, S. Sriboonchitta, and T. Denœux. Prediction of future observations using belief functions: A likelihood-based approach. *International Journal of Approximate Reasoning*, 72:71–94, 2016.

[21] G. J. Klir and B. Yuan. *Fuzzy sets and fuzzy logic. Theory and applications*. Prentice-Hall, Upper Saddle River, NJ., 1995.

[22] E. Miranda, I. Couso, and P. Gil. Random intervals as a model for imprecise information. *Fuzzy Sets and Systems*, 154(3):386–412, 2005.

[23] I. Molchanov. *Theory of Random Sets*. Springer, New York, 2005.

[24] P.-A. Monney. *A Mathematical Theory of Arguments for Statistical Evidence*. Contributions to Statistics. Physica-Verlag, Heidelberg, 2003.

[25] S. Nahmias. Fuzzy variables. *Fuzzy Sets and Systems*, 1(2):97–110, 1978.

[26] H. T. Nguyen. On random sets and belief functions. *Journal of Mathematical Analysis and Applications*, 65:531–542, 1978.

[27] K. B. Petersen and M. S. Pedersen. The matrix cookbook, nov 2012. `http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html`.

[28] M. L. Puri and D. A. Ralescu. Fuzzy random variables. *Journal of Mathematical Analysis and Applications*, 114(2):409–422, 1986.

[29] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.

[30] G. Shafer. Constructive probability. *Synthese*, 48(1):1–60, 1981.

[31] G. Shafer. Belief functions and parametric models (with discussion). *J. Roy. Statist. Soc. Ser. B*, 44:322–352, 1982.

[32] G. Shafer. Dempster's rule of combination. *International Journal of Approximate Reasoning*, 79:26–40, 2016.

[33] P. P. Shenoy. Using possibility theory in expert systems. *Fuzzy Sets and Systems*, 52(2):129–142, 1992.

[34] P. Smets. Belief functions on real numbers. *International Journal of Approximate Reasoning*, 40(3):181–223, 2005.

[35] R. R. Yager. On the normalization of fuzzy belief structure. *International Journal of Approximate Reasoning*, 14:127–153, 1996.

[36] L. A. Zadeh. Fuzzy sets. *Inform. Control*, 8:338–353, 1965.

[37] L. A. Zadeh. Probability measures of fuzzy events. *J. Math. Analysis and Appl.*, 10:421–427, 1968.

[38] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

Appendix A. Proof of Proposition 1

Commutativity is obvious. To prove associativity, let us consider three random sets $(\Omega_i, \sigma_i, P_i, \Theta, \sigma_\Theta, \overline{X}_i)$, $i = 1, 2, 3$. Consider the combined random set

$$(\Omega_1 \times \Omega_2 \times \Omega_3, \sigma_1 \otimes \sigma_2 \otimes \sigma_3, P_{123}, \Theta, \sigma_\Theta, \overline{X}_{1\cap2\cap3}), \tag{A.1}$$

where

$$\overline{X}_{1\cap2\cap3}(\omega_1, \omega_2, \omega_3) = \overline{X}_1(\omega_1) \cap \overline{X}_2(\omega_2) \cap \overline{X}_3(\omega_3),$$

$$P_{123} = (P_1 \times P_2 \times P_3)(\cdot \mid \Theta_{123}^*),$$

and

$$\Theta_{123}^* = \{(\omega_1, \omega_2, \omega_3) \in \Omega_1 \times \Omega_2 \times \Omega_2 : \overline{X}_{1\cap2\cap3}(\omega_1, \omega_2, \omega_3) \neq \emptyset\}.$$

We will show that we get the same result by combining $\overline{X}_1$ with $\overline{X}_2$ first, and then combining the result with $\overline{X}_3$. Combining the first two random sets, we get

$$(\Omega_1 \times \Omega_2, \sigma_1 \otimes \sigma_2, P_{12}, \Theta, \sigma_\Theta, \overline{X}_{1\cap2}),$$

with $\overline{X}_{1\cap2}(\omega_1, \omega_2) = \overline{X}_1(\omega_1) \cap \overline{X}_2(\omega_2)$, $P_{12} = (P_1 \times P_2)(\cdot \mid \Theta_{12}^*)$ and

$$\Theta_{12}^* = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \overline{X}_{1\cap2}(\omega_1, \omega_2) \neq \emptyset\}.$$

Combining it with $\overline{X}_3$ we get

$$(\Omega_1 \times \Omega_2 \times \Omega_3, \sigma_1 \otimes \sigma_2 \otimes \sigma_3, P_{(12)3}, \Theta, \sigma_\Theta, \overline{X}_{1\cap2\cap3}), \tag{A.2}$$

with $P_{(12)3} = (P_{12} \times P_3)(\cdot \mid \Theta_{123}^*)$. Comparing (A.1) and (A.2), we see that we only need to show that $P_{(12)3} = P_{123}$. For any event $C \subseteq \Theta_{123}^*$ and any $\omega_3 \in \Omega_3$, let $C_{\omega_3} = \{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : (\omega_1, \omega_2, \omega_3) \in C\}$. By definition of the product measure $P_{12} \times P_3$ (see [18, page 144]), we have

$$P_{(12)3}(C) = \frac{(P_{12} \times P_3)(C)}{(P_{12} \times P_3)(\Theta_{123}^*)} = \frac{1}{(P_{12} \times P_3)(\Theta_{123}^*)} \int P_{12}(C_{\omega_3}) dP_3(\omega_3) \tag{A.3}$$

Now, as $C \subseteq \Theta_{123}^*$, for any $(\omega_1, \omega_2) \in C_{\omega_3}$, $\overline{X}_1(\omega_1) \cap \overline{X}_2(\omega_2) \neq \emptyset$. Consequently, $C_{\omega_3} \subseteq \Theta_{12}^*$, so

$$P_{12}(C_{\omega_3}) = \frac{(P_1 \times P_2)(C_{\omega_3})}{(P_1 \times P_2)(\Theta_{12}^*)}. \tag{A.4}$$

35

From (A.3) and (A.4), we get

$$P_{(12)3}(C) = \frac{1}{(P_{12} \times P_3)(\Theta_{123}^*)(P_1 \times P_2)(\Theta_{12}^*)} \int (P_1 \times P_2)(C_{\omega_3}) dP_3(\omega_3) \tag{A.5a}$$

$$= \frac{(P_1 \times P_2 \times P_3)(C)}{(P_{12} \times P_3)(\Theta_{123}^*)(P_1 \times P_2)(\Theta_{12}^*)}. \tag{A.5b}$$

Now,

$$P_{123}(C) = \frac{(P_1 \times P_2 \times P_3)(C)}{(P_1 \times P_2 \times P_3)(\Theta_{123}^*)}. \tag{A.6}$$

As $P_{(12)3}(\Theta_{123}^*) = P_{123}(\Theta_{123}^*) = 1$, the denominators in (A.5b) and (A.6) are equal, and $P_{(12)3} = P_{123}$.

Appendix B. Proof of Proposition 4

Commutativity is obvious. To prove associativity, consider three random fuzzy sets

$$(\Omega_i, \sigma_i, P_i, \Theta, \sigma_\Theta, \widetilde{X}_i), \quad i = 1, 2, 3.$$

Let $\widetilde{\Theta}_{12}^*$ be the fuzzy subset of $\Omega_1 \times \Omega_2$ with membership function

$$\widetilde{\Theta}_{12}^*(\omega_1, \omega_2) = \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\right),$$

and let $\widetilde{\Theta}_{(12)3}^*$ and $\widetilde{\Theta}_{123}^*$ be the fuzzy subsets of $\Omega_1 \times \Omega_2 \times \Omega_3$ defined, respectively, as

$$\widetilde{\Theta}_{(12)3}^*(\omega_1, \omega_2, \omega_3) = \mathsf{hgt}\left(\left[\widetilde{X}_1(\omega_1) \odot \widetilde{X}_2(\omega_1)\right]\widetilde{X}_3(\omega_3)\right)$$

and

$$\widetilde{\Theta}_{123}^*(\omega_1, \omega_2, \omega_3) = \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\widetilde{X}_3(\omega_3)\right).$$

Let $\widetilde{P}_{12} = (P_1 \times P_2)(\cdot \mid \widetilde{\Theta}_{12}^*)$, $\widetilde{P}_{(12)3} = (\widetilde{P}_{12} \times P_3)(\cdot \mid \widetilde{\Theta}_{(12)3}^*)$, and $\widetilde{P}_{123} = (P_1 \times P_2 \times P_3)(\cdot \mid \widetilde{\Theta}_{123}^*)$. We only need to show that $\widetilde{P}_{(12)3} = \widetilde{P}_{123}$. For any $B \in \sigma_1 \otimes \sigma_2 \otimes \sigma_3$, we have

$$\widetilde{P}_{(12)3}(B) \propto \int_{\Omega_1 \times \Omega_2} \int_{\Omega_3} B(\omega_1, \omega_2, \omega_3) \mathsf{hgt}\left(\left[\widetilde{X}_1(\omega_1) \odot \widetilde{X}_2(\omega_1)\right]\widetilde{X}_3(\omega_3)\right) dP_3(\omega_3) d\widetilde{P}_{12}(\omega_1, \omega_2)$$

$$\propto \int_{\Omega_1} \int_{\Omega_2} \int_{\Omega_3} B(\omega_1, \omega_2, \omega_3) \mathsf{hgt}\left(\left[\widetilde{X}_1(\omega_1) \odot \widetilde{X}_2(\omega_1)\right]\widetilde{X}_3(\omega_3)\right) \times$$

$$\mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\right) dP_3(\omega_3) dP_2(\omega_2) dP_1(\omega_1).$$

Now,

$$\mathsf{hgt}\left(\left[\widetilde{X}_1(\omega_1) \odot \widetilde{X}_2(\omega_1)\right]\widetilde{X}_3(\omega_3)\right) = \mathsf{hgt}\left(\frac{\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_1)}{\mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2))}\widetilde{X}_3(\omega_3)\right)$$

$$= \frac{\mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\widetilde{X}_3(\omega_3))}{\mathsf{hgt}(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_1))}.$$

36

Hence,

$$\widetilde{P}_{(12)3}(B) \propto \int_{\Omega_1} \int_{\Omega_2} \int_{\Omega_3} B(\omega_1, \omega_2, \omega_3) \mathsf{hgt}\left(\widetilde{X}_1(\omega_1)\widetilde{X}_2(\omega_2)\widetilde{X}_3(\omega_3)\right) dP_3(\omega_3)dP_2(\omega_2)dP_1(\omega_1),$$

which proves that $\widetilde{P}_{(12)3} = \widetilde{P}_{123}$, and the associativity of $\oplus$.

Appendix C. Proof of Proposition 6

We have

$$pl_{\widetilde{X}}(x) = \mathbb{E}_M[\varphi(x; M, h)] \tag{C.1}$$

$$= \int_{-\infty}^{+\infty} \varphi(x; m, h)\phi(m; \mu, \sigma)dm \tag{C.2}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{h}{2}(x-m)^2\right) \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) dm. \tag{C.3}$$

From Proposition 3, the integrand can be written as

$$\exp\left(-\frac{(m-\mu_0)^2}{2\sigma_0^2}\right) \exp\left(-\frac{h(x-\mu)^2}{2(1+h\sigma^2)}\right),$$

with

$$\mu_0 = \frac{xh + \mu/\sigma^2}{h + 1/\sigma^2} = \frac{xh\sigma^2 + \mu}{h\sigma^2 + 1}$$

and

$$\sigma_0 = \sqrt{\frac{1}{h + 1/\sigma^2}} = \frac{\sigma}{\sqrt{1 + h\sigma^2}}.$$

Consequently,

$$pl_{\widetilde{X}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{h(x-\mu)^2}{2(1+h\sigma^2)}\right) \underbrace{\int_{-\infty}^{+\infty} \exp\left(-\frac{(m-\mu_0)^2}{2\sigma_0^2}\right) dm}_{\sigma_0\sqrt{2\pi}} \tag{C.4}$$

$$= \frac{1}{\sqrt{1 + h\sigma^2}} \exp\left(-\frac{h(x-\mu)^2}{2(1+h\sigma^2)}\right). \tag{C.5}$$

Appendix D. Proof of Proposition 7

If $h = 0$, we have, trivially, $Bel_{\widetilde{X}}([x, y]) = 0$ and $Pl_{\widetilde{X}}([x, y]) = 1$ for all $x \leq y$. Let us assume that $h > 0$. We have

$$Pl_{\widetilde{X}}([x, y]) = \mathbb{P}(M \leq x)\mathbb{E}[\varphi(x; M, h) \mid M \leq x] +$$
$$\mathbb{P}(x < M \leq y) \times 1 + \mathbb{P}(M > y)\mathbb{E}[\varphi(y; M, h) \mid M > y], \tag{D.1}$$

which can be written as

$$Pl_{\widetilde{X}}([x,y]) = \Phi\left(\frac{x-\mu}{\sigma}\right)\mathbb{E}[\varphi(x;M,h) \mid M \leq x]+$$
$$\Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right) +$$
$$\left[1 - \Phi\left(\frac{y-\mu}{\sigma}\right)\right]\mathbb{E}[\varphi(y;M,h) \mid M > y]. \quad \text{(D.2)}$$

Conditionally on $M \leq x$, M has a truncated normal distribution on $(-\infty, x]$ with pdf

$$f(m) = \frac{1}{\sigma\sqrt{2\pi}}\frac{\exp\left(\frac{-(m-\mu)^2}{2\sigma^2}\right)}{\Phi\left(\frac{x-\mu}{\sigma}\right)}\mathbf{1}_{(-\infty,x]}(m).$$

Consequently,

$$\mathbb{E}[\varphi(x;M,h) \mid M \leq x] = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{\Phi\left(\frac{x-\mu}{\sigma}\right)}\underbrace{\int_{-\infty}^{x}\exp\left(-\frac{h}{2}(x-m)^2\right)\exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right)dm}_{I}.$$
$$\text{(D.3)}$$

From Proposition 3, integral I in (D.3) can be written as

$$I = \sigma_0\sqrt{2\pi}\Phi\left(\frac{x-\mu_0}{\sigma_0}\right)\exp\left(-\frac{(x-\mu)^2}{2(h^{-1}+\sigma^2)}\right),$$

with

$$\mu_0 = \frac{xh\sigma^2 + \mu}{h\sigma^2 + 1} \quad \text{and} \quad \sigma_0 = \frac{\sigma}{\sqrt{h\sigma^2 + 1}}.$$

Consequently,

$$\mathbb{E}[\varphi(x;M,h) \mid M \leq x] = \frac{1}{\Phi\left(\frac{x-\mu}{\sigma}\right)}pl_{\widetilde{X}}(x)\Phi\left(\frac{x-\mu}{\sigma\sqrt{h\sigma^2+1}}\right).$$

Using similar calculations, we find

$$\mathbb{E}[\varphi(y;M,h) \mid M > y] = \frac{1}{1 - \Phi\left(\frac{y-\mu}{\sigma}\right)}pl_{\widetilde{X}}(y)\left[1 - \Phi\left(\frac{y-\mu}{\sigma\sqrt{h\sigma^2+1}}\right)\right],$$

which concludes the proof of (24).

Now, let us consider (23). We have

$$Bel_{\widetilde{X}}([x,y]) = 1 - Pl_{\widetilde{X}}((-\infty,x] \cup [y,+\infty)),$$

and

$$Pl_{\widetilde{X}}((-\infty, x] \cup [y, +\infty)) = \mathbb{P}(M \le x) \times 1 +$$
$$\mathbb{P}(x < M \le (x+y)/2)\mathbb{E}[\varphi(x; M, h) \mid x < M \le (x+y)/2] +$$
$$\mathbb{P}((x+y)/2 < M \le y)\mathbb{E}[\varphi(y; M, h) \mid (x+y)/2 < M \le y] + \mathbb{P}(M > y) \times 1, \quad \text{(D.4)}$$

which can be written as

$$Pl_{\widetilde{X}}((-\infty, x] \cup [y, +\infty)) = \Phi\left(\frac{x-\mu}{\sigma}\right) +$$
$$\left[\Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]\mathbb{E}[\varphi(x; M, h) \mid x < M \le (x+y)/2] +$$
$$\left[\Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right)\right]\mathbb{E}[\varphi(y; M, h) \mid (x+y)/2 < M \le y] +$$
$$1 - \Phi\left(\frac{y-\mu}{\sigma}\right). \quad \text{(D.5)}$$

Conditionally on $x < M \le (x+y)/2$, M has a truncated normal distribution on $(x, (x+y)/2]$ with pdf

$$f(m) = \frac{1}{\sigma\sqrt{2\pi}} \frac{\exp\left(\frac{-(m-\mu)^2}{2\sigma^2}\right)}{\Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right)} \mathbf{1}_{(x,(x+y)/2]}(m).$$

Consequently,

$$\mathbb{E}[\varphi(x; M, h) \mid x < M \le (x+y)/2] = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{\Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right)} \times$$
$$\underbrace{\int_x^{(x+y)/2} \exp\left(-\frac{h}{2}(x-m)^2\right) \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) dm}_{I'}. \quad \text{(D.6)}$$

The integral in (D.6) is, with the same notations as before,

$$I' = \sigma_0\sqrt{2\pi}\left[\Phi\left(\frac{(x+y)/2-\mu_0}{\sigma_0}\right) - \Phi\left(\frac{x-\mu_0}{\sigma_0}\right)\right]\exp\left(-\frac{(x-\mu)^2}{2(h^{-1}+\sigma^2)}\right).$$

Consequently,

$$\mathbb{E}[\varphi(x; M, h) \mid x < M \le (x+y)/2] =$$
$$\frac{1}{\Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right)} pl_{\widetilde{X}}(x)\left[\Phi\left(\frac{(x+y)/2-\mu+h\sigma^2(y-x)/2}{\sigma\sqrt{h\sigma^2+1}}\right)\right.$$
$$\left. -\Phi\left(\frac{x-\mu}{\sigma\sqrt{h\sigma^2+1}}\right)\right]. \quad \text{(D.7)}$$

39

Similarly, we find

$$\mathbb{E}[\varphi(y; M, h) \mid (x+y)/2 < M \leq y] =$$

$$\frac{1}{\Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(\frac{(x+y)/2-\mu}{\sigma}\right)} pl_{\widetilde{X}}(y) \left[\Phi\left(\frac{y-\mu}{\sigma\sqrt{h\sigma^2+1}}\right) - \right.$$

$$\left. \Phi\left(\frac{(x+y)/2 - \mu - (y-x)h\sigma^2/2}{\sigma\sqrt{h\sigma^2+1}}\right)\right]. \quad \text{(D.8)}$$

The expressions of $Pl_{\widetilde{X}}((-\infty, x] \cup [y, +\infty))$ and $Bel_{\widetilde{X}}([x, y])$ follow.

Appendix E. Proof of Proposition 8

Let $\widetilde{X}(\omega) = \mathsf{GFN}(M(\omega), h)$ be the image of $\omega \in \Omega$ by $\widetilde{X}$, with $M \sim N(\mu, \sigma^2)$. For any $\alpha \in (0, 1]$, its alpha-cut is the random interval

$$^{\alpha}\widetilde{X}(\omega) = \left[M(\omega) - \sqrt{\frac{-2\ln\alpha}{h}}, M(\omega) + \sqrt{\frac{-2\ln\alpha}{h}}\right].$$

Consequently, from (16), the lower and upper expectation of $\widetilde{X}$ are

$$\mathbb{E}_*(\widetilde{X}) = \mu - \int_0^1 \sqrt{\frac{-2\ln\alpha}{h}} d\alpha,$$

and

$$\mathbb{E}^*(\widetilde{X}) = \mu + \int_0^1 \sqrt{\frac{-2\ln\alpha}{h}} d\alpha.$$

By the change of variable $\beta = \sqrt{-2(\ln\alpha)/h}$, we get

$$\int_0^1 \sqrt{\frac{-2\ln\alpha}{h}} d\alpha = h \int_0^{+\infty} \beta^2 \exp\left(-\frac{h\beta^2}{2}\right) d\beta.$$

Now, the second-order moment of the normal distribution $N(0, 1/h)$ is

$$\sqrt{\frac{h}{2\pi}} \int_{-\infty}^{+\infty} \beta^2 \exp\left(-\frac{h\beta^2}{2}\right) d\beta = \frac{1}{h},$$

from which we get

$$h \int_0^{+\infty} \beta^2 \exp\left(-\frac{h\beta^2}{2}\right) d\beta = h \cdot \frac{1}{h} \sqrt{\frac{\pi}{2h}} = \sqrt{\frac{\pi}{2h}}.$$

Appendix F. Proof of Lemma 1

The conditional density of (M_1, M_2) is

$$f(m_1, m_2 \mid \widetilde{F}) = \frac{f(m_1, m_2)\widetilde{F}(m_1, m_2)}{\iint f(m_1, m_2)\widetilde{F}(m_1, m_2)dm_1 dm_2}. \tag{F.1}$$

The numerator on the right-hand side of (F.1) is

$$\frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[\left(\frac{m_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{m_2 - \mu_2}{\sigma_2}\right)^2\right]\right\} \exp\left\{-\frac{\overline{h}(m_1 - m_2)^2}{2}\right\}$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[m_1^2\left(\frac{1}{\sigma_1^2} + \overline{h}\right) - \frac{2m_1\mu_1}{\sigma_1^2} + \frac{\mu_1^2}{\sigma_1^2} + \right.\right.$$
$$\left.\left. m_2^2\left(\frac{1}{\sigma_2^2} + \overline{h}\right) - \frac{2m_2\mu_2}{\sigma_2^2} + \frac{\mu_2^2}{\sigma_2^2} - 2\overline{h}m_1 m_2\right]\right\}. \tag{F.2}$$

Now, the two-dimensional Gaussian density with parameters $(\widetilde{\mu}_1, \widetilde{\mu}_2, \widetilde{\sigma}_1, \widetilde{\sigma}_2, \rho)$ equals

$$\frac{1}{2\pi\widetilde{\sigma}_1\widetilde{\sigma}_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{m_1 - \widetilde{\mu}_1}{\widetilde{\sigma}_1}\right)^2 - \right.\right.$$
$$\left.\left. 2\rho\left(\frac{m_1 - \widetilde{\mu}_1}{\widetilde{\sigma}_1}\right)\left(\frac{m_2 - \widetilde{\mu}_2}{\widetilde{\sigma}_2}\right) + \left(\frac{m_2 - \widetilde{\mu}_2}{\widetilde{\sigma}_2}\right)^2\right]\right\}. \tag{F.3}$$

Equating the second and first-order terms inside the exponentials in (F.2) and (F.3) gives us

$$\widetilde{\sigma}_1 = \frac{1}{1 - \rho^2}\left(\frac{1}{\sigma_1^2} + \overline{h}\right)^{-1} \tag{F.4a}$$

$$\widetilde{\sigma}_2 = \frac{1}{1 - \rho^2}\left(\frac{1}{\sigma_2^2} + \overline{h}\right)^{-1} \tag{F.4b}$$

$$\rho = \frac{\overline{h}\sigma_1\sigma_2}{\sqrt{(1 + \overline{h}\sigma_1^2)(1 + \overline{h}\sigma_2^2)}} \tag{F.4c}$$

$$\widetilde{\mu}_1 = \frac{\mu_1\widetilde{\sigma}_1^2}{\sigma_1^2} + \rho\mu_2\frac{\widetilde{\sigma}_1\widetilde{\sigma}_2}{\sigma_2^2} \tag{F.4d}$$

$$\widetilde{\mu}_2 = \frac{\mu_2\widetilde{\sigma}_2^2}{\sigma_2^2} + \rho\mu_1\frac{\widetilde{\sigma}_1\widetilde{\sigma}_2}{\sigma_1^2}. \tag{F.4e}$$

Replacing ρ by its expression (F.4c) in (F.4a) and (F.4b) yields (28c) and (28d). Replacing ρ, $\widetilde{\sigma}_1$ and $\widetilde{\sigma}_2$ by their expressions in (F.4d) and (F.4e) gives (28a) and (28b).

Finally, the degree of conflict between GRFN's $\widetilde{X}_1 \sim \widetilde{N}(\mu_1, \sigma_1^2, h_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\mu_2, \sigma_2^2, h_2)$ is

$$\kappa = 1 - (P_1 \times P_2)(\widetilde{\Theta}^*),$$

with

$$(P_1 \times P_2)(\widetilde{\Theta}^*) = \iint f(m_1, m_2)\widetilde{F}(m_1, m_2)dm_1dm_2.$$

Taking the ratio of (F.2) to (F.3), we get

$$\iint f(m_1, m_2)\widetilde{F}(m_1, m_2)dm_1dm_2 =$$

$$\frac{\widetilde{\sigma}_1\widetilde{\sigma}_2}{\sigma_1\sigma_2}\sqrt{1 - \rho^2}\exp\left\{-\frac{1}{2}\left[\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right] + \frac{1}{2(1 - \rho^2)}\left[\frac{\widetilde{\mu}_1^2}{\widetilde{\sigma}_1^2} + \frac{\widetilde{\mu}_2^2}{\widetilde{\sigma}_2^2} - 2\rho\frac{\widetilde{\mu}_1\widetilde{\mu}_2}{\widetilde{\sigma}_1\widetilde{\sigma}_2}\right]\right\}.$$

Appendix G. Proof of Proposition 10

From (29), $h_{12} = +\infty$ and the combined GRFN $\widetilde{N}(\widetilde{\mu}_{12}, \widetilde{\sigma}_{12}^2, h_{12})$ is probabilistic. From (30) and (31),

$$\widetilde{\mu}_{12} = \lim_{h_1 \to +\infty} \frac{\widetilde{\mu}_1 + \frac{h_2}{h_1}\widetilde{\mu}_2}{1 + \frac{h_2}{h_1}} = \widetilde{\mu}_1,$$

and

$$\widetilde{\sigma}_{12}^2 = \lim_{h_1 \to +\infty} \frac{\widetilde{\sigma}_1^2 + \frac{h_2^2}{h_1^2}\widetilde{\sigma}_2^2 + 2\rho\frac{h_2}{h_1}\widetilde{\sigma}_1\widetilde{\sigma}_2}{(1 + \frac{h_2}{h_1})^2} = \widetilde{\sigma}_1^2.$$

From (28f),

$$\overline{h} = \lim_{h_1 \to +\infty} \frac{h_2}{1 + \frac{h_2}{h_1}} = h_2.$$

From (28a) and (28c),

$$\widetilde{\mu}_1 = \frac{\mu_1(1 + h_2\sigma_2^2) + \mu_2 h_2\sigma_1^2}{1 + h_2(\sigma_1^2 + \sigma_2^2)},$$

and

$$\widetilde{\sigma}_1^2 = \frac{\sigma_1^2(1 + h_2\sigma_2^2)}{1 + h_2(\sigma_1^2 + \sigma_2^2)}.$$

Now, using Proposition 3, the product of the probability density of X_1 and the contour function of $\widetilde{X}_2$ can be written as

$$f_{X_1}(x)pl_{\widetilde{X}_2}(x) \propto \exp\left(-\frac{1}{2}\frac{(x - \mu_1)^2}{\sigma_1^2}\right)\exp\left(-\frac{h_2(x - \mu_2)^2}{2(1 + h_2\sigma_2^2)}\right)$$

$$\propto \exp\left(-\frac{1}{2\sigma_{12}^2}(x - \mu_{12})^2\right),$$

with

$$\frac{1}{\sigma_{12}^2} = \frac{1}{\sigma_1^2} + \frac{h_2}{1 + h_2\sigma_2^2} = \frac{1 + h_2(\sigma_1^2 + \sigma_2^2)}{\sigma_1^2(1 + h_2\sigma_2^2)}$$

and

$$\mu_{12} = \frac{\frac{1}{\sigma_1^2}\mu_1 + \frac{h_2}{1 + h_2\sigma_2^2}\mu_2}{\frac{1}{\sigma_1^2} + \frac{h_2}{1 + h_2\sigma_2^2}} = \frac{\mu_1(1 + h_2\sigma_2^2) + \mu_2 h_2\sigma_1^2}{1 + h_2(\sigma_1^2 + \sigma_2^2)}.$$

We can check that $\mu_{12} = \widetilde{\mu}_1$ and $\sigma_{12}^2 = \widetilde{\sigma}_1^2$.

Appendix H. Proof of Proposition 12

We have

$$pl_{\widetilde{X}}(\boldsymbol{x}) = \mathbb{E}_M[\varphi(\boldsymbol{x}; \boldsymbol{M}, \boldsymbol{H})] \tag{H.1}$$

$$= \int_{\mathbb{R}^p} \varphi(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{H})\phi(\boldsymbol{m}; \boldsymbol{\mu}, \Sigma)d\boldsymbol{m} \tag{H.2}$$

$$= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \int_{\mathbb{R}^p} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{m})\right) \times \tag{H.3}$$

$$\exp\left(-\frac{1}{2}(\boldsymbol{m} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{m} - \boldsymbol{\mu})\right) d\boldsymbol{m}. \tag{H.4}$$

From Proposition 3, the integrand can be written as

$$\exp\left(-\frac{1}{2}(\boldsymbol{m} - \boldsymbol{\mu}_0)^T\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{m} - \boldsymbol{\mu}_0)\right)\exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T(\boldsymbol{H}^{-1} + \boldsymbol{\Sigma})^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right),$$

with

$$\boldsymbol{\mu}_0 = (\boldsymbol{H} + \boldsymbol{\Sigma}^{-1})^{-1}(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})$$

and

$$\boldsymbol{\Sigma}_0 = (\boldsymbol{H} + \boldsymbol{\Sigma}^{-1})^{-1}.$$

Consequently,

$$pl_{\widetilde{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T(\boldsymbol{H}^{-1} + \boldsymbol{\Sigma})^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right) \times \tag{H.5}$$

$$\underbrace{\int_{\mathbb{R}^p} \exp\left(-\frac{1}{2}(\boldsymbol{m} - \boldsymbol{\mu}_0)^T\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{m} - \boldsymbol{\mu}0)\right) d\boldsymbol{m}}_{(2\pi)^{p/2}|\boldsymbol{\Sigma}_0|^{1/2}} \tag{H.6}$$

$$= \left(\frac{|\boldsymbol{\Sigma}_0|}{|\boldsymbol{\Sigma}|}\right)^{1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T(\boldsymbol{H}^{-1} + \boldsymbol{\Sigma})^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right) \tag{H.7}$$

$$= \frac{1}{|I_p + \boldsymbol{\Sigma}H|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T(\boldsymbol{H}^{-1} + \boldsymbol{\Sigma})^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right). \tag{H.8}$$

Appendix I. Proof of Lemma 2

The conditional density of $\boldsymbol{M} = (\boldsymbol{M}_1, \boldsymbol{M}_2)$ is

$$f(\boldsymbol{m}_1, \boldsymbol{m}_2 \mid \widetilde{F}) = \frac{f(\boldsymbol{m}_1, \boldsymbol{m}_2)\widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2)}{\int_{\mathbb{R}^{2p}} f(\boldsymbol{m}_1, \boldsymbol{m}_2)\widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2)d\boldsymbol{m}_1 d\boldsymbol{m}_2}. \tag{I.1}$$

The numerator on the right-hand side of (I.1) is

$$f(\boldsymbol{m}_1, \boldsymbol{m}_2)\widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) = \phi(\boldsymbol{m}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)\phi(\boldsymbol{m}_2; \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)\times$$

$$\exp\left\{-\frac{1}{2}(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T\overline{\boldsymbol{H}}(\boldsymbol{m}_1 - \boldsymbol{m}_2)\right\}, \quad (I.2)$$

which can be written as

$$f(\boldsymbol{m}_1, \boldsymbol{m}_2)\widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) = \frac{1}{(2\pi)^p |\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2|^{1/2}} \exp\left(-\frac{Z}{2}\right)$$

with

$$Z = \boldsymbol{m}_1^T(\boldsymbol{\Sigma}_1^{-1} + \overline{\boldsymbol{H}})\boldsymbol{m}_1 + \boldsymbol{m}_2^T(\boldsymbol{\Sigma}_2^{-1} + \overline{\boldsymbol{H}})\boldsymbol{m}_2 - 2\boldsymbol{m}_1^T\overline{\boldsymbol{H}}\boldsymbol{m}_2 - 2\boldsymbol{m}_1^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 -$$
$$2\boldsymbol{m}_2^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 + \boldsymbol{\mu}_1^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2. \quad \text{(I.3)}$$

Now, the $2p$-dimensional Gaussian density with mean $\widetilde{\boldsymbol{\mu}}$ and covariance matrix $\widetilde{\boldsymbol{\Sigma}}$ equals

$$\phi(\boldsymbol{m}; \widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}}) = \frac{1}{(2\pi)^p |\widetilde{\boldsymbol{\Sigma}}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{m} - \boldsymbol{\mu})^T \widetilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{m} - \boldsymbol{\mu})\right\}. \quad \text{(I.4)}$$

Decomposing vector $\widetilde{\boldsymbol{\mu}}$ as $\widetilde{\boldsymbol{\mu}} = (\widetilde{\boldsymbol{\mu}}_1, \widetilde{\boldsymbol{\mu}}_2)$, with $\widetilde{\boldsymbol{\mu}}_1, \widetilde{\boldsymbol{\mu}}_2 \in \mathbb{R}^p$, and $\widetilde{\boldsymbol{\Sigma}}^{-1}$ as

$$\widetilde{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{C} \end{pmatrix},$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are $p \times p$ matrices, we can rewrite (I.4) as

$$\phi(\boldsymbol{m}; \widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}}) = \frac{1}{(2\pi)^p |\widetilde{\boldsymbol{\Sigma}}|^{1/2}} \exp\left\{-\frac{1}{2}Z'\right\}$$

with

$$Z' = \boldsymbol{m}_1^T\boldsymbol{A}\boldsymbol{m}_1 - 2\boldsymbol{m}_1^T\boldsymbol{A}\widetilde{\boldsymbol{\mu}}_1 + \widetilde{\boldsymbol{\mu}}_1^T\boldsymbol{A}\widetilde{\boldsymbol{\mu}}_1 + \boldsymbol{m}_2^T\boldsymbol{C}\boldsymbol{m}_2 - 2\boldsymbol{m}_2^T\boldsymbol{C}\widetilde{\boldsymbol{\mu}}_2 + \widetilde{\boldsymbol{\mu}}_2^T\boldsymbol{C}\widetilde{\boldsymbol{\mu}}_u +$$
$$2\boldsymbol{m}_2^T\boldsymbol{B}\boldsymbol{m}_1 - 2\boldsymbol{m}_2^T\boldsymbol{B}\boldsymbol{\mu}_1 - 2\boldsymbol{m}_1^T\boldsymbol{B}\boldsymbol{\mu}_2 + 2\boldsymbol{\mu}_2^T\boldsymbol{B}\boldsymbol{\mu}_1. \quad \text{(I.5)}$$

Equating the second-order terms in (I.3) and (I.5), we get

$$\boldsymbol{A} = \boldsymbol{\Sigma}_1^{-1} + \overline{\boldsymbol{H}}, \quad \boldsymbol{C} = \boldsymbol{\Sigma}_2^{-1} + \overline{\boldsymbol{H}}, \quad \boldsymbol{B} = -\overline{\boldsymbol{H}}.$$

Equating the first-order terms, we get

$$\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 = \boldsymbol{A}\widetilde{\boldsymbol{\mu}}_1 + \boldsymbol{B}\widetilde{\boldsymbol{\mu}}_2 = (\boldsymbol{\Sigma}_1^{-1} + \overline{\boldsymbol{H}})\widetilde{\boldsymbol{\mu}}_1 - \overline{\boldsymbol{H}}\widetilde{\boldsymbol{\mu}}_2, \quad \text{(I.6a)}$$
$$\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 = \boldsymbol{B}\widetilde{\boldsymbol{\mu}}_1 + \boldsymbol{C}\widetilde{\boldsymbol{\mu}}_2 = -\overline{\boldsymbol{H}}\widetilde{\boldsymbol{\mu}}_1 + (\boldsymbol{\Sigma}_2^{-1} + \overline{\boldsymbol{H}})\widetilde{\boldsymbol{\mu}}_2. \quad \text{(I.6b)}$$

Multiplying both sides of (I.6a) and (I.6b) by $\overline{\boldsymbol{H}}^{-1}$, we get

$$(\overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{I}_p)\widetilde{\boldsymbol{\mu}}_1 - \widetilde{\boldsymbol{\mu}}_2 = \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 \quad \text{(I.7)}$$
$$-\widetilde{\boldsymbol{\mu}}_1 + (\overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1} + \boldsymbol{I}_p)\widetilde{\boldsymbol{\mu}}_2 = \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2, \quad \text{(I.8)}$$

which can be written in matrix form

$$
\begin{pmatrix} \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{I}_p & -\boldsymbol{I}_p \\ -\boldsymbol{I}_p & \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1} + \boldsymbol{I}_p \end{pmatrix} \begin{pmatrix} \widetilde{\boldsymbol{\mu}}_1 \\ \widetilde{\boldsymbol{\mu}}_2 \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_1^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \overline{\boldsymbol{H}}^{-1}\boldsymbol{\Sigma}_2^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix},
$$

from which we obtain (36).

Finally, the degree of conflict between GRFV's $\widetilde{X}_1 \sim \widetilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{H}_1)$ and $\widetilde{X}_2 \sim \widetilde{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{H}_2)$ is

$$
\kappa = 1 - (P_1 \times P_2)(\widetilde{\Theta}^*) = 1 - \int_{\mathbb{R}^{2p}} f(\boldsymbol{m}_1, \boldsymbol{m}_2) \widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) d\boldsymbol{m}_1 d\boldsymbol{m}_2.
$$

Taking the ratio of (I.2) to (I.4), we get

$$
\int_{\mathbb{R}^{2p}} f(\boldsymbol{m}_1, \boldsymbol{m}_2) \widetilde{F}(\boldsymbol{m}_1, \boldsymbol{m}_2) d\boldsymbol{m}_1 d\boldsymbol{m}_2 =
$$

$$
\sqrt{\frac{|\widetilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}} \exp\left\{ -\frac{1}{2} \left[\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\mu}_2 - \widetilde{\boldsymbol{\mu}}^T \widetilde{\boldsymbol{\Sigma}}^{-1} \widetilde{\boldsymbol{\mu}} \right] \right\}.
$$

Appendix J. Proof of Lemma 3

The membership function of the projection of fuzzy vector $\mathsf{GFV}(\boldsymbol{m}, \boldsymbol{H})$ on Θ_1 is

$$
\varphi(\boldsymbol{x}_1) = \max_{\boldsymbol{x}_2} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{m}) \right) = \exp\left(-\frac{1}{2} \min_{\boldsymbol{x}_2} Z \right), \tag{J.1}
$$

with $Z = (\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{m})$. Now,

$$
Z = (\boldsymbol{x}_1 - \boldsymbol{m}_1, \boldsymbol{x}_2 - \boldsymbol{m}_2) \begin{pmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 - \boldsymbol{m}_1 \\ \boldsymbol{x}_2 - \boldsymbol{m}_2 \end{pmatrix} \tag{J.2a}
$$

$$
= (\boldsymbol{x}_1 - \boldsymbol{m}_1)^T \boldsymbol{H}_{11}(\boldsymbol{x}_1 - \boldsymbol{m}_1) + (\boldsymbol{x}_2 - \boldsymbol{m}_2)^T \boldsymbol{H}_{21}(\boldsymbol{x}_1 - \boldsymbol{m}_1) + \tag{J.2b}
$$
$$
(\boldsymbol{x}_1 - \boldsymbol{m}_1)^T \boldsymbol{H}_{12}(\boldsymbol{x}_2 - \boldsymbol{m}_2) + (\boldsymbol{x}_2 - \boldsymbol{m}_2)^T \boldsymbol{H}_{22}(\boldsymbol{x}_2 - \boldsymbol{m}_2).
$$

Using $\boldsymbol{H}_{21} = \boldsymbol{H}_{12}^T$, the gradient of Z with respect to $\boldsymbol{x}_2$ can be written as

$$
\frac{\partial Z}{\partial \boldsymbol{x}_2} = 2\boldsymbol{H}_{21}(\boldsymbol{x}_1 - \boldsymbol{m}_1) + 2\boldsymbol{H}_{22}(\boldsymbol{x}_2 - \boldsymbol{m}_2).
$$

Setting $\frac{\partial Z}{\partial \boldsymbol{x}_2} = 0$, and assuming $\boldsymbol{H}_{22}$ to be nonsingular, we get

$$
(\boldsymbol{x}_2 - \boldsymbol{m}_2) = -\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}(\boldsymbol{x}_1 - \boldsymbol{m}_1). \tag{J.3}
$$

Replacing $(\boldsymbol{x}_2 - \boldsymbol{m}_2)$ by its expression (J.3) in (J.2) and using (J.1), we finally get

$$
\varphi(\boldsymbol{x}_1) = \exp\left(-\frac{1}{2}(\boldsymbol{x}_1 - \boldsymbol{m}_1)^T \boldsymbol{H}_{11}'(\boldsymbol{x}_1 - \boldsymbol{m}_1) \right),
$$

with

$$
\boldsymbol{H}_{11}' = \boldsymbol{H}_{11} - \boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}.
$$

45