



HAL
open science

Parametric families of continuous belief functions based on generalized Gaussian random fuzzy numbers

Thierry Denœux

► **To cite this version:**

Thierry Denœux. Parametric families of continuous belief functions based on generalized Gaussian random fuzzy numbers. *Fuzzy Sets and Systems*, 2023, 471, pp.108679. 10.1016/j.fss.2023.108679 . hal-04060251v2

HAL Id: hal-04060251

<https://hal.science/hal-04060251v2>

Submitted on 6 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Parametric families of continuous belief functions based on generalized Gaussian random fuzzy numbers

Thierry Denœux^{a,b}

^a*Université de technologie de Compiègne, CNRS
UMR 7253 Heudiasyc, Compiègne, France*

^b*Institut universitaire de France, Paris, France*

Abstract

The theory of epistemic random fuzzy sets is a general theory of uncertainty encompassing both possibility theory and the Dempster-Shafer theory of belief functions as special cases. Within this framework, Gaussian random fuzzy numbers have recently been introduced as a practical model of uncertainty about real variables. However, the limited flexibility of this model does not allow it to represent all kinds of beliefs encountered in applications. In this paper, it is extended in two ways. First, we study one-to-one transformations of random fuzzy sets and show that such transformations commute with combination. This property allows us to define parametric families of easily combinable random fuzzy numbers and vectors on different frames based on the Gaussian model. We then go one step further by studying mixtures of random fuzzy variables, which provide a very flexible model making it possible to construct belief functions on continuous frames with arbitrary complexity. To demonstrate the applicability and practical interest of these models, two applications are studied: the elicitation of expert beliefs about numerical quantities, and generalized Bayesian inference with weak prior information represented by random fuzzy numbers.

Keywords: Belief functions, evidence theory, possibility theory, random fuzzy sets, uncertainty, statistical inference, elicitation.

1. Introduction

The Dempster-Shafer (DS) theory of belief functions [5, 26, 12] and possibility theory [30, 15, 13] are two powerful frameworks for representing partial information and reasoning with uncertainty. Whereas DS theory makes it possible to represent partially reliable evidence, possibility theory allows us to express uncertainty associated with vague information such as conveyed by fuzzy sets. In [8, 14], we have argued that DS and possibility theories can be viewed as two specializations of a more general theory of “epistemic random fuzzy sets”. A random fuzzy set (RFS), also called “random fuzzy variable”, maps each element of a probability space to a fuzzy subset of a set Θ . It is, thus, a model of evidence that can

Email address: Thierry.Denoeux@utc.fr (Thierry Denœux)

10 be both uncertain and fuzzy. In this framework, a possibility distribution can be viewed
11 as a constant RFS, while a random set (a notion underlying DS theory) corresponds to the
12 special case where all images are crisp. Random fuzzy sets induced by independent pieces of
13 evidence can be combined by the (generalized) product-intersection rule, which generalizes
14 both Dempster’s rule of combination and the normalized product intersection of possibility
15 distributions.

16 Whereas the theory of belief functions has been defined from the start in a very general
17 setting [27], most applications have used only belief functions on finite spaces. This limita-
18 tion was mainly due to the absence of general enough parametric families of belief functions
19 in continuous spaces that could easily be defined and combined by Dempster’s rule of com-
20 bination. In [9], we have proposed Gaussian Random Fuzzy Numbers (GRFNs) as a model
21 for defining belief functions on the real line. A GRFN is a “doubly Gaussian model”: it
22 can be seen either as a Gaussian possibility distribution whose mode is a Gaussian random
23 variable, or as a Gaussian probability distribution whose mean is a Gaussian fuzzy set. We
24 have also proposed Gaussian Random Fuzzy Vectors (GRFVs) as a multidimensional exten-
25 sion of GRFNs, which makes it possible to construct belief functions in \mathbb{R}^p for $p \geq 1$. The
26 families of GRFNs and GRFVs are closed under the product-intersection operation, which
27 makes them suitable for evidential reasoning with continuous variables. An application to
28 machine learning was presented in [9, 11].

29 Practical as it may be, the GRFN model is quite restricted. The domain of a GRFN is
30 the whole real line, making the model unsuitable for representing belief functions on a real
31 interval such as $(0, +\infty)$ or $[a, b]$. Furthermore, a GRFN is unimodal (the contour function
32 has a unique maximum) and symmetric about the mean μ , i.e., intervals of the form $[\mu - r, \mu]$
33 and $[\mu, \mu + r]$, for any $r > 0$, have the same degree of belief; these properties may not always
34 reflect an agent’s actual beliefs. The GRFV model is also inadequate for representing, e.g.,
35 beliefs on probabilities or proportions, for which the domain of interest is the probability
36 simplex. It is thus of interest to define more flexible parameterized families of random fuzzy
37 numbers and vectors with different supports and different “shapes”, while maintaining the
38 closure property under the product-intersection rule.

39 In this paper¹ the objective stated above is achieved in two ways. We first study *bijective*
40 *transformations* of RFSs and show that such transformations commute with combination,
41 i.e., applying the transformation before or after the combination yields the same result. This
42 property is exploited to define easily combinable random fuzzy numbers and vectors based
43 on GRFNs and GRFVs. We then go one step further by studying *mixtures* of RFSs, which
44 provide a very flexible model making it possible to construct belief functions on \mathbb{R}^p with
45 virtually unlimited complexity. Finally, we combine the two ideas and propose *mixtures of*
46 *transformed GRFNs and GRFVs* as a very general model of RFSs (and associated belief
47 functions) easily combinable using the product-intersection rule.

48 To demonstrate the applicability and practical interest of the flexible models introduced
49 in this paper, two applications will be discussed. The first one concerns the *elicitation*

¹This paper is an extended version of the short paper [10] to be presented at the 2023 IEEE International Conference on Fuzzy Systems, Songdo Incheon, Korea, August 13-17, 2023.

50 of expert beliefs about a numerical quantity of interest. Without getting into the delicate
51 methodological problems posed by real elicitation experiments, we will demonstrate how
52 the parameters of a random fuzzy number can be fitted to a small number of expert plau-
53 sibility statements. The second application is related to *generalized Bayesian inference*. As
54 discussed in [8, 14], the relative likelihood function can be seen as possibility distribution
55 about the parameter of interest. This sample information can be combined with prior knowl-
56 edge expressed as a RFS, which constitutes a form of “weak prior”, ranging continuously
57 from precise knowledge represented by a probability distribution, to complete ignorance. In
58 many cases, the likelihood function (sometimes after transforming the parameter) is well ap-
59 proximated by a Gaussian possibility distribution and can easily be combined with a GRFN
60 or a mixture thereof to compute a posterior RFS.

61 The rest of this paper is organized as follows. The main definitions related to RFSs and
62 GRFNs are first recalled in Section 2. Transformations of RFSs and mixture models are
63 studied, respectively, in Section 3 and 4. The applications are then discussed in Section 5.
64 Finally, Section 6 concludes the paper.

65 2. Random fuzzy sets

66 To make the paper self-contained, the RFS setting and its relation with belief functions
67 will first be briefly reviewed in Section 2.1. The GRFN and GRFV models will then be
68 recalled, respectively, in Sections 2.2 and 2.3.

69 2.1. General definitions and results

70 *Definition.* Let us consider a probability space $(\Omega, \Sigma_\Omega, P)$, a measurable space (Θ, Σ_Θ) , and
71 a mapping \tilde{X} from Ω to the set $[0, 1]^\Theta$ of fuzzy subsets of Θ (see Figure 1). For any $\alpha \in [0, 1]$,
72 let ${}^\alpha\tilde{X}$ be the mapping from Ω to 2^Θ such that

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

73 where ${}^\alpha[\tilde{X}(\omega)] = \{\theta \in \Theta : \tilde{X}(\omega)(\theta) \geq \alpha\}$ is the weak α -cut of $\tilde{X}(\omega)$. If, for any $\alpha \in [0, 1]$,
74 ${}^\alpha\tilde{X}$ is $\Sigma_\Omega - \Sigma_\Theta$ strongly measurable [24], the tuple $(\Omega, \Sigma_\Omega, P, \Theta, \Sigma_\Theta, \tilde{X})$ is said to be a *random*
75 *fuzzy set* (also called a *fuzzy random variable*) [4]. The images $\tilde{X}(\omega)$ of elements of Ω by
76 \tilde{X} are called the (*fuzzy*) *focal sets* of \tilde{X} . We define the *support* of \tilde{X} as the union of the
77 supports of its focal sets, i.e.,

$$\text{supp}(\tilde{X}) = \bigcup_{\omega \in \Omega} \{\theta \in \Theta : \tilde{X}(\omega)(\theta) > 0\}.$$

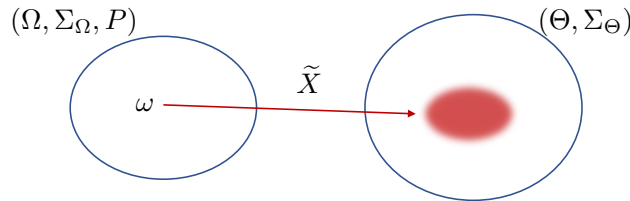


Figure 1: Definition of a random fuzzy set.

78 If Θ is equal to \mathbb{R} or a real interval, and if the images $\tilde{X}(\omega)$ are fuzzy numbers (i.e., normal
79 and convex fuzzy subsets of \mathbb{R}), \tilde{X} is said to be a *random fuzzy number* (RFN). For any
80 $\alpha \in [0, 1]$, the mapping ${}^\alpha\tilde{X}$ is, then, a random interval.

81 *Interpretation.* In epistemic random fuzzy set (ERFS) theory, RFSs are used to represent
82 unreliable and fuzzy evidence: the set Ω is then seen as a *set of interpretations* of a piece of
83 evidence about a variable θ taking values in Θ . If interpretation $\omega \in \Omega$ holds, we know that
84 “ θ is $\tilde{X}(\omega)$ ”, i.e., θ is constrained by the possibility distribution defined by fuzzy set $\tilde{X}(\omega)$,
85 which represents the *meaning* of the evidence if interpretation ω holds. We do not know
86 which interpretation is true, but we hold beliefs about the true interpretation, expressed by
87 the probability measure P on Ω . Such a RFS represents a piece of evidence with uncertain
88 meaning or, equivalently, a *state of knowledge* about some variable θ , hence the adjective
89 “epistemic”. This model should not be confused with alternative interpretations of RFSs as
90 describing a fuzzy data generation mechanism [25, 18], or as imprecise information about a
91 “true” random variable [21, 3].

92 As a toy example illustrating the notion of unreliable and fuzzy evidence, assume that θ
93 represents John’s height in centimeters, with domain $\Theta = [0, 200]$. We receive a testimony
94 telling us that “John is tall”, the linguistic term “tall” being represented by a fuzzy subset
95 \tilde{T} of Θ . Let us denote by $\Omega = \{R, \neg R\}$ the set of interpretations of the evidence, where R
96 stands for “reliable” and $\neg R$ for “not reliable”, and assume that we have 80% confidence
97 that this testimony is reliable, i.e., $P(R) = 0.8$ and $P(\neg R) = 0.2$. If the testimony is reliable,
98 John’s height is constrained by fuzzy set \tilde{T} ; if it is not, we know nothing about John’s height;
99 the mapping \tilde{X} is, thus, defined by $\tilde{X}(R) = \tilde{T}$ and $\tilde{X}(\neg R) = \Theta$.

100 It is clear that this model can encode different kinds of uncertainty: probabilistic uncer-
101 tainty represented by P , and imprecision and/or fuzziness represented by mapping \tilde{X} . If
102 each image $\tilde{X}(\omega)$ is a singleton, only probabilistic uncertainty is present. Pure imprecision
103 or fuzziness corresponds to the case where \tilde{X} is a constant mapping, i.e., there exists some
104 crisp or fuzzy subset \tilde{F} of Θ such that for all $\omega \in \Omega$, $\tilde{X}(\omega) = \tilde{F}$. If \tilde{F} is a crisp subset, no
105 fuzziness is present.

106 *Belief and plausibility functions.* Just as a random set, a RFS induces a belief function,
107 which can be seen as quantifying one’s beliefs based on the available evidence. From now
108 on, we will assume any RFS \tilde{X} to verify the following normalization conditions:

- 109 1. For all $\omega \in \Omega$, the height of $\tilde{X}(\omega)$, defined as $\text{hgt}(\tilde{X}(\omega)) = \sup_{\theta \in \Theta} \tilde{X}(\omega)(\theta)$ is either 0
110 or 1, i.e., $\tilde{X}(\omega)$ is either the empty set, or a normal fuzzy set;
- 111 2. The image $\tilde{X}(\omega)$ is almost surely nonempty, i.e., $P(\{\omega \in \Omega : \tilde{X}(\omega) = \emptyset\}) = 0$.

112 For any $\omega \in \Omega$, let $\Pi_{\tilde{X}(\omega)}$ be the possibility measure on Θ quantifying our beliefs on θ
113 given that interpretation ω holds; it is defined, for any $B \in \Sigma_\Theta$, as

$$\Pi_{\tilde{X}(\omega)}(B) = \sup_{\theta \in B} \tilde{X}(\omega)(\theta). \quad (1a)$$

114 The dual necessity measure $N_{\tilde{X}(\omega)}$ is

$$N_{\tilde{X}(\omega)}(B) = \begin{cases} 1 - \Pi_{\tilde{X}(\omega)}(B^c) & \text{if } \tilde{X}(\omega) \neq \emptyset \\ 0 & \text{otherwise,} \end{cases} \quad (1b)$$

115 where B^c denotes the complement of B . For any $B \in \Sigma_{\Theta}$, let $Bel_{\tilde{X}}(B)$ and $Pl_{\tilde{X}}(B)$ denote,
116 respectively, the *expected necessity* and the *expected possibility* of B :

$$Bel_{\tilde{X}}(B) = \int_{\Omega} N_{\tilde{X}(\omega)}(B) dP(\omega), \quad (2a)$$

117

$$Pl_{\tilde{X}}(B) = \int_{\Omega} \Pi_{\tilde{X}(\omega)}(B) dP(\omega) = 1 - Bel_{\tilde{X}}(B^c). \quad (2b)$$

118 The mappings $B \mapsto Bel_{\tilde{X}}(B)$ and $B \mapsto Pl_{\tilde{X}}(B)$, are, respectively, belief and plausibility
119 functions [31, 4].

120 *Lower and upper expectations of a RFN.* Let \tilde{X} be a RFN, and let F_* and F^* be its lower
121 and upper cumulative distribution functions (cdfs) defined, respectively, as

$$F_*(x) = Bel_{\tilde{X}}((-\infty, x]) \quad \text{and} \quad F^*(x) = Pl_{\tilde{X}}((-\infty, x])$$

122 for all $x \in \mathbb{R}$. We define the *lower and upper expectations* of \tilde{X} (see [5]) as, respectively, the
123 integrals

$$\mathbb{E}_*(\tilde{X}) = \int x dF_*(x) \quad \text{and} \quad \mathbb{E}^*(\tilde{X}) = \int x dF^*(x).$$

124 It can be shown [4, 14] that these integrals can be computed as the means of the lower and
125 upper expectations of the α -cuts of \tilde{X} , i.e.,

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

126 If, for any $\omega \in \Omega$ and any $\alpha \in (0, 1]$, the α -cut $\alpha \tilde{X}(\omega)$ is an interval with lower bound $\alpha \tilde{X}^-(\omega)$
127 and upper bound $\alpha \tilde{X}^+(\omega)$, the lower and upper expectations of $\alpha \tilde{X}$ are, respectively, the
128 expectations of these lower and upper bounds. The lower and upper expectations of \tilde{X} can
129 then be computed by averaging these expectations with respect to α , i.e.,

$$\mathbb{E}_*(\tilde{X}) = \int_0^1 \mathbb{E}(\alpha \tilde{X}^-) d\alpha \quad \text{and} \quad \mathbb{E}^*(\tilde{X}) = \int_0^1 \mathbb{E}(\alpha \tilde{X}^+) d\alpha. \quad (3)$$

130

131 *Combination of RFSs.* The combination of independent pieces of evidence by Dempster's
 132 rule [26] is a key component of DS theory. In possibility theory, conjunctive combination
 133 operators are based on t-norms [16]. In ERFs theory, the product-intersection rule intro-
 134 duced in [8, 14] extends these operators to the general case where evidence is represented
 135 by RFSs.

136 Let us assume that we have a piece of evidence about θ represented by RFS

$$(\Omega_1, \Sigma_1, P_1, \Theta, \Sigma_\Theta, \tilde{X}_1),$$

137 and we receive a second one represented by RFS $(\Omega_2, \Sigma_2, P_2, \Theta, \Sigma_\Theta, \tilde{X}_2)$. We consider the
 138 joint measurable space $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$, where $\Sigma_1 \otimes \Sigma_2$ is the tensor product of σ -algebras
 139 Σ_1 and Σ_2 , and we denote by P the probability measure representing our beliefs about the
 140 interpretations of both items of evidence, *before considering the conflicts between mappings*
 141 \tilde{X}_1 and \tilde{X}_2 . Naturally, we have $P(A \times \Omega_2) = P_1(A)$ for all $A \in \Sigma_1$ and $P(\Omega_1 \times B) = P_2(B)$
 142 for all $B \in \Sigma_2$. The two pieces of evidence are said to be *independent*² iff the following two
 143 conditions hold:

- 144 1. The meaning of one piece of evidence depends only on its own set of interpretations.
 145 That is to say, if we denote by $\tilde{X}_1(\omega_1, \omega_2)$ the meaning of the first piece of evidence
 146 given that the pair of interpretations $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$ holds, we have

$$\forall(\omega_2, \omega'_2) \in \Omega_2^2, \quad \tilde{X}_1(\omega_1, \omega_2) = \tilde{X}_1(\omega_1, \omega'_2)$$

147 and, symmetrically, denoting by $\tilde{X}_2(\omega_1, \omega_2)$ the meaning of the second piece of evidence
 148 given $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$,

$$\forall(\omega_1, \omega'_1) \in \Omega_1^2, \quad \tilde{X}_2(\omega_1, \omega_2) = \tilde{X}_2(\omega'_1, \omega_2).$$

- 149 2. For any $A \in \Sigma_1$ and $B \in \Sigma_2$, our degree of belief that the true interpretation of the
 150 first evidence belongs to A and the true interpretation of the second evidence belongs
 151 to B is $P(A \times B) = P_1(A)P_2(B)$, i.e., P is the product measure $P_1 \times P_2$.

152 Under these two assumptions, we construct a RFS representing the aggregation of the two
 153 pieces of evidence by proceeding in two steps.

154 **Step 1:** We define a new mapping from the product space $\Omega_1 \times \Omega_2$ to $[0, 1]^\Theta$ that maps
 155 each pair of interpretations (ω_1, ω_2) to a fuzzy intersection of $\tilde{X}_1(\omega_1)$ and $\tilde{X}_2(\omega_2)$.
 156 As argued in [16, 8], the normalized product intersection \odot is the most suitable for
 157 combining fuzzy information from independent sources. Furthermore, this operation
 158 is associative. We thus define the following mapping,

$$(\omega_1, \omega_2) \mapsto \tilde{X}_\odot(\omega_1, \omega_2) = \tilde{X}_1(\omega_1) \odot \tilde{X}_2(\omega_2)$$

²The notion of independence defined here concerns items of evidence, and not RFSs. A formal definition of independence for RFSs could be useful, e.g., in relation with the interpretation of RFSs as models of mechanisms for generating fuzzy data; it is outside the scope of this paper.

159

with

$$\left(\tilde{X}_1(\omega_1) \odot \tilde{X}_2(\omega_2)\right)(\theta) = \begin{cases} \frac{\tilde{X}_1(\omega_1)(\theta) \cdot \tilde{X}_2(\omega_2)(\theta)}{\text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2))} & \text{if } \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)) > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

160

We assume that each α -cut ${}^\alpha\tilde{X}_\odot$ is $\Sigma_1 \otimes \Sigma_2$ - Σ_Θ strongly measurable.

161

Step 2: We then need to take into account the conflict between mappings \tilde{X}_1 and \tilde{X}_2 . Pair of interpretations (ω_1, ω_2) such that $\text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)) = 0$ are obviously inconsistent and they must be eliminated. However, in the fuzzy setting, we also need to consider *partially inconsistent* pairs of interpretation such that $0 < \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)) < 1$. This can be achieved by *soft normalization* proposed in [8, 14], which consists in conditioning the product probability measure $P_1 \times P_2$ by the fuzzy subset $\tilde{\Theta}^*$ of consistent pairs of interpretations, with membership function

162

163

164

165

166

167

$$\tilde{\Theta}^*(\omega_1, \omega_2) = \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)). \quad (5)$$

The resulting conditional probability measure $\tilde{P}_{12} = (P_1 \times P_2)(\cdot | \tilde{\Theta}^*)$ has the following expression, for any $C \in \Sigma_1 \otimes \Sigma_2$:

$$\begin{aligned} \tilde{P}_{12}(C) &= \frac{(P_1 \times P_2)(C \cap \tilde{\Theta}^*)}{(P_1 \times P_2)(\tilde{\Theta}^*)} \\ &= \frac{\int_{\Omega_1} \int_{\Omega_2} C(\omega_1, \omega_2) \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)}{\int_{\Omega_1} \int_{\Omega_2} \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)}, \end{aligned} \quad (6)$$

168

where $C(\cdot, \cdot)$ denotes the indicator function of C .

169

The combined RFS, denoted by $\tilde{X}_1 \oplus \tilde{X}_2$ and called the *orthogonal sum* of \tilde{X}_1 and \tilde{X}_2 is, thus, formally defined by the following tuple:

170

$$(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \tilde{P}_{12}, \Theta, \Sigma_\Theta, \tilde{X}_\odot).$$

171

The operator \oplus is commutative and associative; it generalizes both Dempster's rule and the normalized product intersection of possibility distributions.

172

173

2.2. Gaussian random fuzzy sets

174

The important role played by the Gaussian distribution in probability theory and statistics is partly due to the fact that it is amenable to easy calculation. Until recently, such a practical model was missing in DS theory, which hindered its application to uncertain reasoning with real variables. The GRFN model fills this gap by blending Gaussian possibility distributions and Gaussian random variables.

175

176

177

178

179 *Gaussian fuzzy number.* Let us start by recalling the definition of a *Gaussian Fuzzy Number*
 180 (GFN) as a fuzzy subset of \mathbb{R} with membership function

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

181 where $m \in \mathbb{R}$ is the *mode* and $h \in [0, +\infty]$ is the *precision*. Such a fuzzy number will be
 182 denoted by $\text{GFN}(m, h)$. GFNs are easily combined by the normalized product intersection op-
 183 erator (4), as the following property holds: $\text{GFN}(m_1, h_1) \odot \text{GFN}(m_2, h_2) = \text{GFN}(m_{12}, h_1 + h_2)$,
 184 with $m_{12} = (h_1 m_1 + h_2 m_2)/(h_1 + h_2)$. Furthermore, the height of $\text{GFN}(m_1, h_1) \cdot \text{GFN}(m_2, h_2)$
 185 is

$$\text{hgt}[\text{GFN}(m_1, h_1) \cdot \text{GFN}(m_2, h_2)] = \exp\left(-\frac{h_1 h_2 (m_1 - m_2)^2}{2(h_1 + h_2)}\right). \quad (7)$$

186
 187 *Gaussian random fuzzy number.* Let us now consider a probability space $(\Omega, \Sigma_\Omega, P)$ and a
 188 Gaussian random variable (GRV) $M : \Omega \rightarrow \mathbb{R}$ with mean μ and variance σ^2 . The random
 189 fuzzy set $\tilde{X} : \Omega \rightarrow [0, 1]^\mathbb{R}$ defined as

$$\tilde{X}(\omega) = \text{GFN}(M(\omega), h)$$

190 is called a *Gaussian random fuzzy number* (GRFN) with mean μ , variance σ^2 and precision
 191 h , which we write $\tilde{X} \sim \tilde{N}(\mu, \sigma^2, h)$. A GRFN can, thus, be seen as a GFN whose mode is
 192 uncertain and described by a Gaussian probability distribution. It is defined by a location
 193 parameter μ , and two parameters h and σ^2 corresponding, respectively, to possibilistic and
 194 probabilistic uncertainty. In the special case where the precision is infinite, \tilde{X} becomes
 195 equivalent to a GRV with mean μ and variance σ^2 , which we can write: $\tilde{N}(\mu, \sigma^2, +\infty) =$
 196 $N(\mu, \sigma^2)$. If $\sigma^2 = 0$, M is constant and \tilde{X} is equivalent to possibility distribution $\text{GFN}(\mu, h)$,
 197 i.e., $\tilde{N}(\mu, 0, h) = \text{GFN}(\mu, h)$. Finally, when $h = 0$, we have $\tilde{X}(\omega)(x) = 1$ for all $\omega \in \Omega$ and
 198 all $x \in \mathbb{R}$: such a RFS represents total ignorance and the corresponding belief function is
 199 said to be *vacuous*.

200 *Belief and plausibility.* Formulas to compute the plausibility and belief degrees of any real
 201 interval $[x, y]$ induced by a GRFN $\tilde{X} \sim \tilde{N}(\mu, \sigma^2, h)$ are given in [14]. In particular, the
 202 contour function of \tilde{X} is given by

$$pl_{\tilde{X}}(x) = \frac{1}{\sqrt{1 + h\sigma^2}} \exp\left(-\frac{h(x - \mu)^2}{2(1 + h\sigma^2)}\right). \quad (8)$$

203 The lower and upper cdfs have the following expressions:

$$\text{Bel}_{\tilde{X}}((-\infty, x]) = \Phi\left(\frac{x - \mu}{\sigma}\right) - pl_{\tilde{X}}(x) \Phi\left(\frac{x - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right), \quad (9a)$$

204 where Φ is the standard normal cdf, and

$$Pl_{\tilde{X}}((-\infty, x]) = \text{Bel}_{\tilde{X}}((-\infty, x]) + pl_{\tilde{X}}(x). \quad (9b)$$

205 *Lower and upper expectations.* Let $\tilde{X} \sim \tilde{N}(\mu, \sigma^2, h)$ be a GRFN with $h > 0$. As shown in
 206 [14], its lower and upper expectations are, respectively,

$$\mathbb{E}_*(\tilde{X}) = \mu - \sqrt{\frac{\pi}{2h}} \quad \text{and} \quad \mathbb{E}^*(\tilde{X}) = \mu + \sqrt{\frac{\pi}{2h}}. \quad (10)$$

207 *Combination of GRFNs.* Most importantly, as shown in [14], the family of GRFNs is
 208 closed under the product-intersection rule: given two GRFNs $\tilde{X}_1 \sim \tilde{N}(\mu_1, \sigma_1^2, h_1)$ and
 209 $\tilde{X}_2 \sim \tilde{N}(\mu_2, \sigma_2^2, h_2)$, we have $\tilde{X}_1 \oplus \tilde{X}_2 \sim \tilde{N}(\tilde{\mu}_{12}, \tilde{\sigma}_{12}^2, h_1 + h_2)$, with

$$\tilde{\mu}_{12} = \frac{h_1\tilde{\mu}_1 + h_2\tilde{\mu}_2}{h_1 + h_2}, \quad \tilde{\sigma}_{12}^2 = \frac{h_1^2\tilde{\sigma}_1^2 + h_2^2\tilde{\sigma}_2^2 + 2\rho h_1 h_2 \tilde{\sigma}_1 \tilde{\sigma}_2}{(h_1 + h_2)^2}, \quad (11a)$$

210 where

$$\tilde{\mu}_1 = \frac{\mu_1(1 + \bar{h}\sigma_2^2) + \mu_2\bar{h}\sigma_1^2}{1 + \bar{h}(\sigma_1^2 + \sigma_2^2)}, \quad (11b)$$

$$\tilde{\mu}_2 = \frac{\mu_2(1 + \bar{h}\sigma_1^2) + \mu_1\bar{h}\sigma_2^2}{1 + \bar{h}(\sigma_1^2 + \sigma_2^2)}, \quad (11c)$$

$$\tilde{\sigma}_1^2 = \frac{\sigma_1^2(1 + \bar{h}\sigma_2^2)}{1 + \bar{h}(\sigma_1^2 + \sigma_2^2)}, \quad \tilde{\sigma}_2^2 = \frac{\sigma_2^2(1 + \bar{h}\sigma_1^2)}{1 + \bar{h}(\sigma_1^2 + \sigma_2^2)}, \quad (11d)$$

$$\rho = \frac{\bar{h}\sigma_1\sigma_2}{\sqrt{(1 + \bar{h}\sigma_1^2)(1 + \bar{h}\sigma_2^2)}}, \quad (11e)$$

214 and $\bar{h} = h_1 h_2 / (h_1 + h_2)$. The degree of conflict between \tilde{X}_1 and \tilde{X}_2 is given by the following
 215 proposition.

Proposition 1. *The degree of conflict between $\tilde{X}_1 \sim \tilde{N}(\mu_1, \sigma_1^2, h_1)$ and $\tilde{X}_2 \sim \tilde{N}(\mu_2, \sigma_2^2, h_2)$ is*

$$\kappa = \begin{cases} 1 - \frac{\tilde{\sigma}_1 \tilde{\sigma}_2}{\sigma_1 \sigma_2} \sqrt{1 - \rho^2} \exp \left\{ -\frac{1}{2} \left[\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} \right] + \frac{1}{2(1-\rho^2)} \left[\frac{\tilde{\mu}_1^2}{\tilde{\sigma}_1^2} + \frac{\tilde{\mu}_2^2}{\tilde{\sigma}_2^2} - 2\rho \frac{\tilde{\mu}_1 \tilde{\mu}_2}{\tilde{\sigma}_1 \tilde{\sigma}_2} \right] \right\} & \text{if } \sigma_1, \sigma_2 > 0 \\ 1 - \frac{1}{\sqrt{1 + \bar{h}\tilde{\sigma}_1^2}} \exp \left(-\frac{\bar{h}}{2(1 + \bar{h}\tilde{\sigma}_1^2)} (\tilde{\mu}_1 - \mu_2)^2 \right) & \text{if } \sigma_1 \geq 0, \sigma_2 = 0, \end{cases}$$

216 where $\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \rho$ are given by (11).

217 *Proof.* The formula for the case $\sigma_1 > 0, \sigma_2 > 0$ is proved in [14]. The case $\sigma_2 = 0$ can be
 218 treated by replacing $\tilde{\mu}_1, \tilde{\mu}_2, \tilde{\sigma}_1^2, \tilde{\sigma}_2^2$, by their expressions given by (11), and letting σ_2 tend
 219 to zero. \square

2.3. Gaussian random fuzzy vectors

220 Gaussian random fuzzy vectors (GRFVs) are multidimensional extensions of GRFNs:
 221 they are defined as Gaussian fuzzy vectors (GFVs), whose modes are random vectors with
 222 a multidimensional Gaussian distribution. We start by giving the definition of GFVs, after
 223 which we recall that of GRFVs as well as some properties.
 224

225 *Gaussian fuzzy vectors.* A p -dimensional GFV with mode $\mathbf{m} \in \mathbb{R}^p$ and $p \times p$ symmetric
 226 and positive semidefinite precision matrix \mathbf{H} is defined as the normalized fuzzy subset of
 227 \mathbb{R}^p with membership function

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

228 It is denoted as $\widetilde{\text{GFV}}(\mathbf{m}, \mathbf{H})$. The normalized product of two GFV's is still a GFV; more
 229 precisely, the following equality holds: $\widetilde{\text{GFV}}(\mathbf{m}_1, \mathbf{H}_1) \odot \widetilde{\text{GFV}}(\mathbf{m}_2, \mathbf{H}_2) = \widetilde{\text{GFV}}(\mathbf{m}_{12}, \mathbf{H}_{12})$,
 230 with $\mathbf{m}_{12} = (\mathbf{H}_1 + \mathbf{H}_2)^{-1}(\mathbf{H}_1 \mathbf{m}_1 + \mathbf{H}_2 \mathbf{m}_2)$ and $\mathbf{H}_{12} = \mathbf{H}_1 + \mathbf{H}_2$.

231 *Gaussian random fuzzy vectors.* Let $(\Omega, \Sigma_\Omega, P)$ be a probability space, $\mathbf{M} : \Omega \rightarrow \mathbb{R}^p$ a p -
 232 dimensional Gaussian random vector with mean $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Sigma}$, and \mathbf{H} a $p \times p$
 233 symmetric and positive semidefinite real matrix. The random fuzzy set $\widetilde{X} : \Omega \rightarrow [0, 1]^{\mathbb{R}^p}$
 234 defined as $\widetilde{X}(\omega) = \widetilde{\text{GFV}}(\mathbf{M}(\omega), \mathbf{H})$ is called a *Gaussian random fuzzy vector (GRFV)*, which
 235 we denote as $\widetilde{X} \sim \widetilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{H})$.

236 As shown in [14], the contour function of a GRFV is given by the following equation,
 237 which generalizes (8):

$$pl_{\widetilde{X}}(\mathbf{x}) = \frac{1}{|\mathbf{I}_p + \boldsymbol{\Sigma} \mathbf{H}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T (\mathbf{H}^{-1} + \boldsymbol{\Sigma})^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (12)$$

238 where \mathbf{I}_p is the p -dimensional identity matrix.

239 The orthogonal sum of two GRFVs is still a GRFV. Formulas for the mean, variance
 240 matrix and precision matrix of the combined GRFV were derived in [14]. For completeness,
 241 they are recalled in Appendix A.

242 3. Transformations of Gaussian random fuzzy variables

243 As mentioned in Section 1, the GRFN and GRFV models are very convenient for un-
 244 certain reasoning with real variables due to their closure property with respect to the \oplus
 245 operator, but they also have several limitations. In particular, the support of a GRFN is
 246 the whole real line, making it unsuitable for representing evidence about variables taking
 247 values in a strict subset of \mathbb{R} . In this section, we overcome this limitation by considering
 248 bijective transformations of RFSs. The main result, stated in Section 3.1, is that the image
 249 of the orthogonal sum of two RFSs by a bijective mapping is the orthogonal sum of the
 250 images. Some useful transformations of GRFNs are studied in Section 3.2, and a particular
 251 transformation of GRFVs is considered in Section 3.3.

252 3.1. Transformation of a random fuzzy set

253 Let $(\Omega, \Sigma_\Omega, P, \Theta, \Sigma_\Theta, \widetilde{X})$ be a RFS, and $\psi : \Theta \rightarrow \Lambda$ a one-to-one mapping from Θ to
 254 some set Λ . Zadeh's extension principle [29] allows us to extend mapping ψ to fuzzy subsets
 255 of Θ ; specifically, we can define a mapping $\widetilde{\psi} : [0, 1]^\Theta \rightarrow [0, 1]^\Lambda$ such that

$$\forall \widetilde{F} \in [0, 1]^\Theta, \quad \widetilde{\psi}(\widetilde{F})(\lambda) = \sup_{\lambda=\psi(\theta)} \widetilde{F}(\theta) = \widetilde{F}(\psi^{-1}(\lambda)). \quad (13)$$

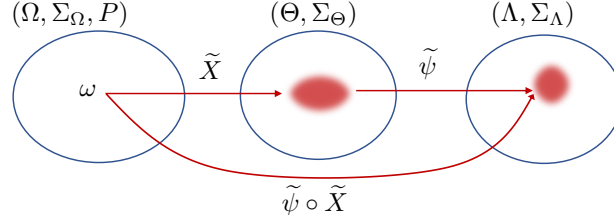


Figure 2: Transformation of a random fuzzy set \tilde{X} by a one-to-one mapping $\psi : \Theta \rightarrow \Lambda$.

We note that mapping $\tilde{\psi}$ is also one-to-one, and its inverse is the extension of ψ^{-1} : for any $\tilde{F} \in [0, 1]^\Theta$,

$$\begin{aligned} (\widetilde{\psi^{-1} \circ \tilde{\psi}})(\tilde{F})(\theta) &= \widetilde{\psi^{-1}(\tilde{\psi}(\tilde{F}))}(\theta) \\ &= \sup_{\theta=\psi^{-1}(\lambda)} \tilde{\psi}(\tilde{F})(\lambda) \\ &= \sup_{\theta=\psi^{-1}(\lambda)} \tilde{F}(\psi^{-1}(\lambda)) = \tilde{F}(\theta). \end{aligned}$$

256 Similarly, $(\tilde{\psi} \circ \widetilde{\psi^{-1}})(\tilde{F}) = \tilde{F}$, so we can write $\widetilde{\psi^{-1}} = \tilde{\psi}^{-1}$.

257 We now consider the composed mapping $\tilde{\psi} \circ \tilde{X}$ from Ω to $[0, 1]^\Lambda$, such that $(\tilde{\psi} \circ \tilde{X})(\omega) =$
 258 $\tilde{\psi}[\tilde{X}(\omega)]$ (see Figure 2). To show that it is a RFS, we start by the following lemma.

259 **Lemma 1.** *The set Σ_Λ containing the images of all elements of Σ_Θ by ψ ,*

$$\Sigma_\Lambda = \{\psi(B) : B \in \Sigma_\Theta\},$$

260 *is a σ -algebra on Λ .*

261 *Proof.* Since Σ_Θ is a σ -algebra, it contains the empty set \emptyset ; consequently, $\psi(\emptyset) = \emptyset \in \Sigma_\Lambda$.
 262 Now, for any $A \in \Sigma_\Lambda$, $B = \psi^{-1}(A) \in \Sigma_\Theta$; hence, $\psi(\Theta \setminus B) = \Lambda \setminus \psi(B) \in \Sigma_\Lambda$. Finally, let
 263 (A_i) , $i \in I$ be a collection of elements of Σ_Λ , and $B_i = \psi^{-1}(A_i)$, $i \in I$ their inverse images.
 264 We have

$$\bigcup_{i \in I} A_i = \bigcup_{i \in I} \psi(B_i) = \psi \left(\bigcup_{i \in I} B_i \right) \in \Sigma_\Lambda.$$

265 □

266 **Proposition 2.** *Let $(\Omega, \Sigma_\Omega, P, \Theta, \Sigma_\Theta, \tilde{X})$ be a RFS, $\psi : \Theta \rightarrow \Lambda$ a one-to-one mapping from*
 267 *Θ to Λ , and $\Sigma_\Lambda = \{\psi(B) : B \in \Sigma_\Theta\}$. The tuple $(\Omega, \Sigma_\Omega, P, \Lambda, \Sigma_\Lambda, \tilde{\psi} \circ \tilde{X})$ is a RFS.*

268 *Proof.* We need to prove that, for any $\alpha \in [0, 1]$ the mapping ${}^\alpha(\tilde{\psi} \circ \tilde{X})$ is $\Sigma_\Omega - \Sigma_\Lambda$ strongly
 269 measurable [24], i.e., for any $A \in \Sigma_\Lambda$,

$$A^* = \{\omega \in \Omega : {}^\alpha(\tilde{\psi} \circ \tilde{X})(\omega) \cap A \neq \emptyset\} \in \Sigma_\Omega.$$

Now,

$$\begin{aligned}
{}^\alpha(\tilde{\psi} \circ \tilde{X})(\omega) &= \{\lambda \in \Lambda : (\tilde{\psi} \circ \tilde{X})(\omega)(\lambda) \geq \alpha\} \\
&= \{\lambda \in \Lambda : \tilde{X}(\omega)[\psi^{-1}(\lambda)] \geq \alpha\} \\
&= \{\lambda \in \Lambda : \psi^{-1}(\lambda) \in {}^\alpha\tilde{X}(\omega)\} \\
&= \psi[{}^\alpha\tilde{X}(\omega)].
\end{aligned}$$

Consequently,

$$\begin{aligned}
A^* &= \{\omega \in \Omega : \psi[{}^\alpha\tilde{X}(\omega)] \cap A \neq \emptyset\} \\
&= \{\omega \in \Omega : \psi[{}^\alpha\tilde{X}(\omega) \cap \psi^{-1}(A)] \neq \emptyset\} \\
&= \{\omega \in \Omega : {}^\alpha\tilde{X}(\omega) \cap \psi^{-1}(A) \neq \emptyset\}.
\end{aligned}$$

270 As ${}^\alpha\tilde{X}$ is $\Sigma_\Omega - \Sigma_\Theta$ strongly measurable and $\psi^{-1}(A) \in \Sigma_\Theta$, it follows that $A^* \in \Sigma_\Omega$. \square

Example 1. Let $\Omega = \{\omega_1, \omega_2\}$, P the probability measure on Ω defined by $P(\{\omega_1\}) = 0.7$, $P(\{\omega_2\}) = 0.3$, and $\tilde{X} : \Omega \rightarrow [0, 1]^\mathbb{R}$ a random fuzzy number with two fuzzy focal sets defined by $\tilde{X}(\omega_1) = \tilde{F}_1 \sim \text{GFN}(1, 1)$ and $\tilde{X}(\omega_2) = \tilde{F}_2 \sim \text{GFN}(2, 0.5)$ (see Figure 3a). Let ψ the bijection from $\Theta = \mathbb{R}$ to $\Lambda = \mathbb{R}_+$ defined by $\psi : \theta \mapsto \exp(\theta)$. From (13), the images of \tilde{F}_1 and \tilde{F}_2 by $\tilde{\psi}$ are, respectively, the following fuzzy subsets of \mathbb{R}_+ :

$$\tilde{\psi}(\tilde{F}_1)(\lambda) = \tilde{F}_1(\log \lambda) = \exp\left(-\frac{1}{2}(\log \lambda - 1)^2\right)$$

and

$$\tilde{\psi}(\tilde{F}_2)(\lambda) = \tilde{F}_2(\log \lambda) = \exp\left(-\frac{1}{4}(\log \lambda - 2)^2\right).$$

271 They are plotted in Figure 3b. Let $\tilde{Y} = \tilde{\psi} \circ \tilde{X}$ the transformation of \tilde{X} by $\tilde{\psi}$; it is defined by
272 $\tilde{Y}(\omega_1) = \tilde{\psi}(\tilde{F}_1)$ and $\tilde{Y}(\omega_2) = \tilde{\psi}(\tilde{F}_2)$.

273 *Belief and plausibility.* Interestingly, the belief and plausibility functions induced by the
274 transformed RFS $\tilde{\psi} \circ \tilde{X}$ have a simple expression in terms of corresponding functions induced
275 by \tilde{X} , as expressed by the following theorem.

276 **Theorem 1.** Let $(\Omega, \Sigma_\Omega, P, \Theta, \Sigma_\Theta, \tilde{X})$ be a RFS, $\psi : \Theta \rightarrow \Lambda$ a one-to-one mapping from Θ
277 to Λ , $\Sigma_\Lambda = \{\psi(B) : B \in \Sigma_\Theta\}$, and $\tilde{\psi} \circ \tilde{X}$ the RFS resulting from the transformation of \tilde{X}
278 by ψ . For any $C \in \Sigma_\Lambda$,

$$Bel_{\tilde{\psi} \circ \tilde{X}}(C) = Bel_{\tilde{X}}(\psi^{-1}(C)), \quad (14a)$$

279 and

$$Pl_{\tilde{\psi} \circ \tilde{X}}(C) = Pl_{\tilde{X}}(\psi^{-1}(C)). \quad (14b)$$

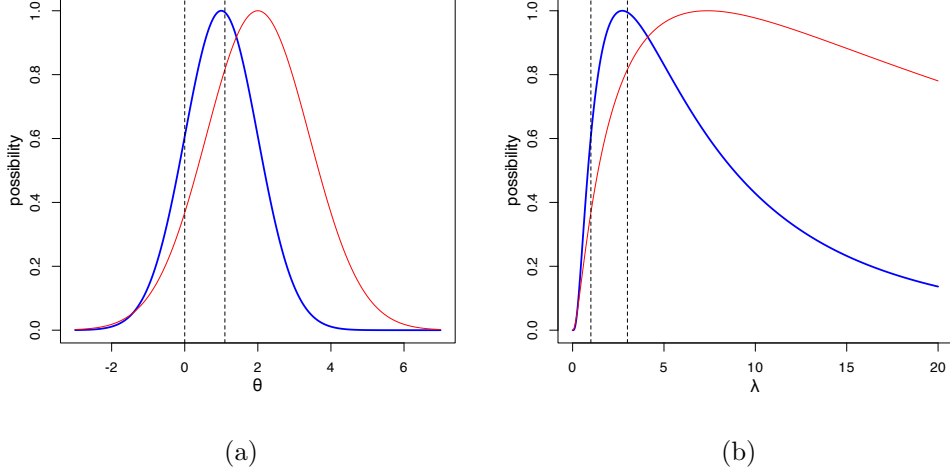


Figure 3: Focal fuzzy sets of \tilde{X} (left) and $\tilde{Y} = \tilde{\psi} \circ \tilde{X}$ (right) in Example 1. The vertical broken lines indicate the intervals $C = [1, 3]$ (right) and $\log(C) = [0, \log(3)]$ (left).

Proof. From (1), for all $C \in \Sigma_\Lambda$,

$$\begin{aligned}
\Pi_{(\tilde{\psi} \circ \tilde{X})(\omega)}(C) &= \sup_{\lambda \in C} (\tilde{\psi} \circ \tilde{X})(\omega)(\lambda) \\
&= \sup_{\lambda \in C} \tilde{X}(\omega)(\psi^{-1}(\lambda)) \\
&= \sup_{\theta \in \psi^{-1}(C)} \tilde{X}(\omega)(\theta) = \Pi_{\tilde{X}(\omega)}(\psi^{-1}(C)),
\end{aligned}$$

280 and, similarly,

$$N_{(\tilde{\psi} \circ \tilde{X})(\omega)}(C) = N_{\tilde{X}(\omega)}(\psi^{-1}(C)).$$

281 The result follows directly using the definition of belief and plausibility function in (2). \square

282 **Example 2.** Continuing Example 1, let $C = [1, 3]$. We can compute $Pl_{\tilde{Y}}(C)$ and $Bel_{\tilde{Y}}(C)$
283 in two ways (see Figure 3):

1. Using the definitions (2):

$$\begin{aligned}
Pl_{\tilde{Y}}(C) &= 0.7 \sup_{\lambda \in [1,3]} \tilde{\psi}(\tilde{F}_1)(\lambda) + 0.3 \sup_{\lambda \in [1,3]} \tilde{\psi}(\tilde{F}_2)(\lambda) \\
&= 0.7 \times 1 + 0.3 \times \exp\left(-\frac{1}{4}(\log(3) - 2)^2\right) \approx 0.94 \\
Bel_{\tilde{Y}}(C) &= 1 - \left(0.7 \sup_{\lambda \notin [1,3]} \tilde{\psi}(\tilde{F}_1)(\lambda) + 0.3 \sup_{\lambda \notin [1,3]} \tilde{\psi}(\tilde{F}_2)(\lambda)\right) \\
&= 1 - \left(0.7 \times \exp\left(-\frac{1}{2}(\log(3) - 1)^2\right) + 0.3 \times 1\right) \approx 0.0034.
\end{aligned}$$

2. Using Theorem 1:

$$\begin{aligned}
Pl_{\tilde{Y}}(C) &= Pl_{\tilde{X}}(\psi^{-1}(C)) \\
&= 0.7 \sup_{\theta \in [0, \log(3)]} \tilde{F}_1(\theta) + 0.3 \sup_{\theta \in [0, \log(3)]} \tilde{F}_2(\theta) \\
&= 0.7 \times 1 + 0.3 \times \exp\left(-\frac{1}{4}(\log 3 - 2)^2\right) \approx 0.94 \\
Bel_{\tilde{Y}}(C) &= Bel_{\tilde{X}}(\psi^{-1}(C)) \\
&= 1 - \left(0.7 \sup_{\theta \notin [0, \log(3)]} \tilde{F}_1(\theta) + 0.3 \sup_{\theta \notin [0, \log(3)]} \tilde{F}_2(\theta)\right) \\
&= 1 - \left(0.7 \times \exp\left(-\frac{1}{2}(\log 3 - 1)^2\right) + 0.3 \times 1\right) \approx 0.0034.
\end{aligned}$$

284 *Combination.* Let now consider the combination of two transformed RFSs $\psi \circ \tilde{X}_1$ and $\psi \circ \tilde{X}_2$
285 with the same transformation ψ . The following lemma states that the image of the product
286 intersection of two fuzzy subsets of Θ is equal to the product intersection of their images, and
287 the *degree of conflict* (defined as the height of the product intersection before normalization)
288 of the fuzzy subsets equals that of their images.

289 **Lemma 2.** Let \tilde{F} and \tilde{G} be two fuzzy subsets of Θ . We have

$$\tilde{\psi}(\tilde{F} \odot \tilde{G}) = \tilde{\psi}(\tilde{F}) \odot \tilde{\psi}(\tilde{G})$$

290 and

$$hgt(\tilde{\psi}(\tilde{F}) \cdot \tilde{\psi}(\tilde{G})) = hgt(\tilde{F} \cdot \tilde{G}).$$

Proof. For any $\lambda \in \Lambda$,

$$\tilde{\psi}(\tilde{F} \odot \tilde{G})(\lambda) = (\tilde{F} \odot \tilde{G})[\psi^{-1}(\lambda)] \tag{15a}$$

$$= \frac{\tilde{F}[\psi^{-1}(\lambda)]\tilde{G}[\psi^{-1}(\lambda)]}{\sup_{\lambda'} \tilde{F}[\psi^{-1}(\lambda')]\tilde{G}[\psi^{-1}(\lambda')]} \tag{15b}$$

$$= \frac{\tilde{\psi}(\tilde{F})(\lambda)\tilde{\psi}(\tilde{G})(\lambda)}{\sup_{\lambda'} \tilde{\psi}(\tilde{F})(\lambda')\tilde{\psi}(\tilde{G})(\lambda')} \tag{15c}$$

$$= (\tilde{\psi}(\tilde{F}) \odot \tilde{\psi}(\tilde{G}))(\lambda). \tag{15d}$$

291 Now, the degree of conflict between $\tilde{\psi}(\tilde{F})$ and $\tilde{\psi}(\tilde{G})$ is the denominator on the right-hand
292 side of (15b). It is equal to

$$\sup_{\lambda \in \Lambda} \tilde{F}[\psi^{-1}(\lambda)]\tilde{G}[\psi^{-1}(\lambda)] = \sup_{\theta \in \Theta} \tilde{F}(\theta)\tilde{G}(\theta).$$

293

□

294 We can now state the main result of this section.

295 **Theorem 2.** Let $(\Omega_i, \Sigma_i, P_i, \Theta, \Sigma_\Theta, \tilde{X}_i)$, $i = 1, 2$, be two RFSs representing independent
 296 evidence. We have

$$\tilde{\psi} \circ (\tilde{X}_1 \oplus \tilde{X}_2) = (\tilde{\psi} \circ \tilde{X}_1) \oplus (\tilde{\psi} \circ \tilde{X}_2).$$

297 *Proof.* As recalled in Section 2.1, the orthogonal sum of $\tilde{\psi} \circ \tilde{X}_1$ and $\tilde{\psi} \circ \tilde{X}_2$ is defined by
 298 mapping

$$(\omega_1, \omega_2) \mapsto (\tilde{\psi} \circ \tilde{X}_1)(\omega_1) \odot (\tilde{\psi} \circ \tilde{X}_2)(\omega_2),$$

299 and the joint probability measure $P_1 \times P_2$ conditioned by the fuzzy subset of $\Omega_1 \times \Omega_2$ with
 300 membership function

$$\Theta^*(\omega_1, \omega_2) = \text{hgt} \left((\tilde{\psi} \circ \tilde{X}_1)(\omega_1) \cdot (\tilde{\psi} \circ \tilde{X}_2)(\omega_2) \right).$$

301 Now, from Lemma 2,

$$(\tilde{\psi} \circ \tilde{X}_1)(\omega_1) \odot (\tilde{\psi} \circ \tilde{X}_2)(\omega_2) = \tilde{\psi} \left[\tilde{X}_1(\omega_1) \odot \tilde{X}_2(\omega_2) \right].$$

302 Hence, the mappings from Ω to $[0, 1]^\Lambda$ associated to $\tilde{\psi} \circ (\tilde{X}_1 \oplus \tilde{X}_2)$ and $(\tilde{\psi} \circ \tilde{X}_1) \oplus (\tilde{\psi} \circ \tilde{X}_2)$
 303 are identical. Furthermore, from Lemma 2,

$$\text{hgt}((\tilde{\psi} \circ \tilde{X}_1)(\omega_1) \cdot (\tilde{\psi} \circ \tilde{X}_2)(\omega_2)) = \text{hgt}(\tilde{X}_1(\omega_1) \cdot \tilde{X}_2(\omega_2)).$$

304 The fuzzy conditioning event in $[0, 1]^{\Omega_1 \times \Omega_2}$ associated to $\tilde{X}_1 \oplus \tilde{X}_2$ and $(\tilde{\psi} \circ \tilde{X}_1) \oplus (\tilde{\psi} \circ \tilde{X}_2)$
 305 are, thus, also identical, which completes the proof. \square

306 **Example 3.** Continuing Examples 1 and 2, let us consider another RFS defined by $\Omega' =$
 307 $\{\omega'_1, \omega'_2\}$, $P'(\{\omega'_1\}) = 0.6$, $P'(\{\omega'_2\}) = 0.4$, $\tilde{X}'(\omega'_1) = \tilde{F}'_1 \sim \text{GFN}(0, 1)$, $\tilde{X}'(\omega'_2) = \tilde{F}'_2 \sim$
 308 $\text{GFN}(1.5, 1)$. Let $\tilde{Y} = \tilde{\psi} \circ \tilde{X}$ and $\tilde{Y}' = \tilde{\psi} \circ \tilde{X}'$ and assume that we want to compute their
 309 orthogonal sum $\tilde{Y} \oplus \tilde{Y}'$. From Theorem 2, we can first combine \tilde{X} and \tilde{X}' , and then compose
 310 the result with $\tilde{\psi}$. Using (7), we obtain

$$\text{hgt}(\tilde{F}_1 \cdot \tilde{F}'_1) = 0.78, \quad \text{hgt}(\tilde{F}_1 \cdot \tilde{F}'_2) = 0.94$$

311

$$\text{hgt}(\tilde{F}_2 \cdot \tilde{F}'_1) = 0.51, \quad \text{hgt}(\tilde{F}_2 \cdot \tilde{F}'_2) = 0.96$$

312 Consequently, from (6), the joint probability measure on $\Omega \times \Omega'$ conditioned on fuzzy subset
 313 $\tilde{\Theta}$ defined by (5) is

$$P''(\{(\omega_1, \omega'_1)\}) \propto 0.7 \times 0.6 \times 0.78$$

314

$$P''(\{(\omega_1, \omega'_2)\}) \propto 0.7 \times 0.4 \times 0.94$$

315

$$P''(\{(\omega_2, \omega'_1)\}) \propto 0.3 \times 0.6 \times 0.51$$

316

$$P''(\{(\omega_2, \omega'_2)\}) \propto 0.3 \times 0.4 \times 0.96.$$

317 After normalization, we get

$$P''(\{(\omega_1, \omega'_1)\}) \approx 0.41, \quad P''(\{(\omega_1, \omega'_2)\}) \approx 0.33$$

318

$$P''(\{(\omega_2, \omega'_1)\}) \approx 0.12, \quad P''(\{(\omega_2, \omega'_2)\}) \approx 0.14.$$

Now, the combined mapping is $\tilde{X} \oplus \tilde{X}'$ is

$$(\tilde{X} \oplus \tilde{X}')(\omega_1, \omega'_1) = \tilde{F}_1 \odot \tilde{F}'_1 \sim \text{GFN}(0.5, 2)$$

$$(\tilde{X} \oplus \tilde{X}')(\omega_1, \omega'_2) = \tilde{F}_1 \odot \tilde{F}'_2 \sim \text{GFN}(1.25, 2)$$

$$(\tilde{X} \oplus \tilde{X}')(\omega_2, \omega'_1) = \tilde{F}_2 \odot \tilde{F}'_1 \sim \text{GFN}(2/3, 1.5)$$

$$(\tilde{X} \oplus \tilde{X}')(\omega_2, \omega'_2) = \tilde{F}_2 \odot \tilde{F}'_2 \sim \text{GFN}(5/3, 1.5).$$

319 Finally, we have

$$(\tilde{Y} \oplus \tilde{Y}')(\omega_1, \omega'_1) = \tilde{\psi}(\tilde{F}_1 \odot \tilde{F}'_1), \quad (\tilde{Y} \oplus \tilde{Y}')(\omega_1, \omega'_2) = \tilde{\psi}(\tilde{F}_1 \odot \tilde{F}'_2)$$

320

$$(\tilde{Y} \oplus \tilde{Y}')(\omega_2, \omega'_1) = \tilde{\psi}(\tilde{F}_2 \odot \tilde{F}'_1), \quad (\tilde{Y} \oplus \tilde{Y}')(\omega_2, \omega'_2) = \tilde{\psi}(\tilde{F}_2 \odot \tilde{F}'_2).$$

321 3.2. Transformed Gaussian Random Fuzzy Numbers

322 Applying the idea developed in Section 3.1 to GRFNs makes it possible to define a wide
 323 variety of parametric families of random fuzzy numbers and associated belief functions on
 324 the real line. Let $\tilde{X} \sim \tilde{N}(\mu, \sigma^2, h)$ be a GRFN, and ψ a one-to-one mapping from \mathbb{R} to
 325 $\Lambda \subseteq \mathbb{R}$. Let $\tilde{\psi} \circ \tilde{X}$ be the result of the transformation of \tilde{X} by ψ . We will say that $\tilde{\psi} \circ \tilde{X}$ is
 326 a *transformed GRFN* (or t-GRFN) and we will write $\tilde{\psi} \circ \tilde{X} \sim T\tilde{N}(\mu, \sigma^2, h, \psi^{-1})$. For any
 327 random fuzzy number \tilde{Y} , it is clear that

$$\tilde{Y} \sim T\tilde{N}(\mu, \sigma^2, h, \psi^{-1}) \Leftrightarrow \tilde{\psi}^{-1} \circ \tilde{Y} \sim \tilde{N}(\mu, \sigma^2, h). \quad (16)$$

328 From Theorem 2, given two t-GRFNs $\tilde{Y}_i \sim T\tilde{N}(\mu_i, \sigma_i^2, h_i, \psi^{-1})$, $i = 1, 2$, we have $Y_1 \oplus Y_2 \sim$
 329 $T\tilde{N}(\tilde{\mu}_{12}, \tilde{\sigma}_{21}^2, h_1 + h_2, \psi^{-1})$, where $\tilde{\mu}_{12}$ and $\tilde{\sigma}_{21}^2$ are given by (11).

330 Hereafter, we will consider three cases for the choice of function ψ allowing us to define
 331 belief functions on the positive real line, on a closed real interval, or on the whole real line.

332 *Lognormal random fuzzy numbers.* Using a one-to-one mapping from \mathbb{R} to $(0, +\infty)$ allows
 333 us to define a random fuzzy number with support equal to the positive real line. Choosing
 334 $\psi = \exp$, we obtain a *lognormal random fuzzy number (RFN)* $\tilde{Y} \sim T\tilde{N}(\mu, \sigma^2, h, \log)$. From
 335 (16), $\tilde{Y} \sim T\tilde{N}(\mu, \sigma^2, h, \log)$ if and only if $\log(\tilde{Y}) \sim \tilde{N}(\mu, \sigma^2, h)$. A lognormal random variable
 336 is recovered when $h = +\infty$. From (8) and (14b), the contour function of \tilde{Y} is

$$pl_{\tilde{Y}}(y) = \frac{1}{\sqrt{1+h\sigma^2}} \exp\left(-\frac{h(\log y - \mu)^2}{2(1+h\sigma^2)}\right).$$

Similarly, the lower and upper cdfs of \tilde{Y} can easily be computed from (9) and (14) as

$$Bel_{\tilde{Y}}((-\infty, y]) = \Phi\left(\frac{\log y - \mu}{\sigma}\right) - pl_{\tilde{Y}}(y)\Phi\left(\frac{\log y - \mu}{\sigma\sqrt{h\sigma^2 + 1}}\right),$$

and $Pl_{\tilde{Y}}((-\infty, y]) = Bel_{\tilde{Y}}((-\infty, y]) + pl_{\tilde{Y}}(y)$.

Expressions for the lower and upper expectations of a lognormal RFN are given in the following proposition.

Proposition 3. *The lower and upper expectation of a lognormal RFN $\tilde{Y} \sim T\tilde{N}(\mu, \sigma^2, h, \log)$ are given, respectively, by*

$$\mathbb{E}_*(\tilde{Y}) = \sqrt{2\pi} \exp\left(\mu + \frac{\sigma^2}{2} + \frac{1}{2h}\right) \left[\phi\left(\frac{1}{\sqrt{h}}\right) - \frac{1}{\sqrt{h}} \left(1 - \Phi\left(\frac{1}{\sqrt{h}}\right)\right) \right] \quad (17a)$$

and

$$\mathbb{E}^*(\tilde{Y}) = \sqrt{2\pi} \exp\left(\mu + \frac{\sigma^2}{2} + \frac{1}{2h}\right) \left[\phi\left(\frac{1}{\sqrt{h}}\right) + \frac{1}{\sqrt{h}} \Phi\left(\frac{1}{\sqrt{h}}\right) \right], \quad (17b)$$

where ϕ is the standard normal probability density function.

Proof. See Appendix B □

We note that the expectation of a lognormal random variable is recovered in the limit when h tends to infinity, as

$$\lim_{h \rightarrow +\infty} \mathbb{E}_*(\tilde{Y}) = \lim_{h \rightarrow +\infty} \mathbb{E}^*(\tilde{Y}) = \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

Example 4. *Figure 4 displays two lognormal RFNs*

$$\tilde{Y}_1 \sim T\tilde{N}(1, 1, 5, \log) \quad \text{and} \quad \tilde{Y}_2 \sim T\tilde{N}(2, 0.1, 2, \log),$$

as well as their orthogonal sum $\tilde{Y}_1 \oplus \tilde{Y}_2$. For each RFN, we plot ten realizations, the contour function, the lower and upper expectations, as well as the lower and upper cdfs.

Logit-normal random fuzzy numbers. Any cdf F can be used to define a RFN with support equal to interval $[0, 1]$ (or more generally, using an additional affine transformation, an interval $[a, b]$ with $b > a$). A natural choice is the cdf of the standard logistic distribution, $F_L(x) = [1 + \exp(-x)]^{-1}$. The corresponding quantile function is the logit function,

$$F_L^{-1}(y) = \text{logit}(y) = \log \frac{y}{1-y}.$$

A RFN $\tilde{Y} \sim T\tilde{N}(\mu, \sigma, h, \text{logit})$ will be said to be *logit-normal*. A logit-normal random variable $[1]$ is recovered when $h = +\infty$. From (8) and (14b), the contour function of \tilde{Y} is

$$pl_{\tilde{Y}}(y) = \frac{1}{\sqrt{1+h\sigma^2}} \exp\left(-\frac{h(\text{logit}(y) - \mu)^2}{2(1+h\sigma^2)}\right).$$

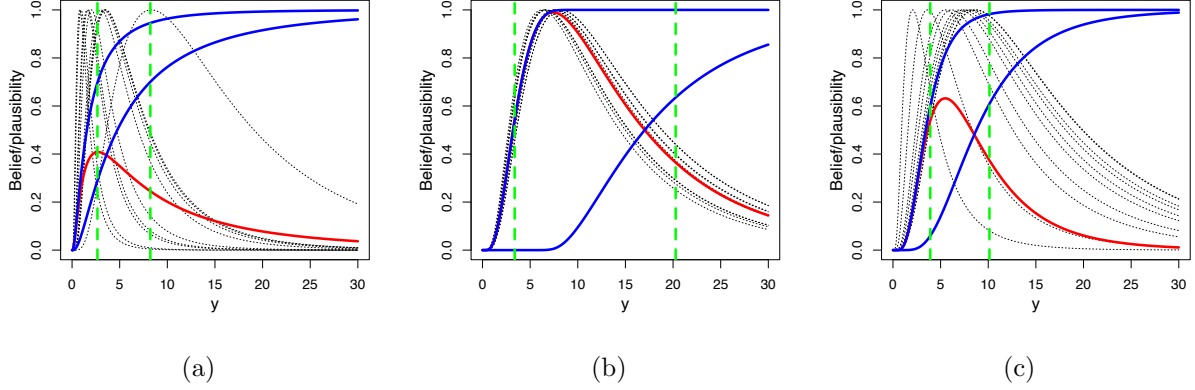


Figure 4: (a) and (b): From left to right, two lognormal random fuzzy numbers $\tilde{Y}_1 \sim T\tilde{N}(1, 1, 5, \log)$ and $\tilde{Y}_2 \sim T\tilde{N}(2, 0.1, 2, \log)$; (c): combined lognormal random fuzzy number $\tilde{Y}_1 \oplus \tilde{Y}_2$. For each RFN, we plot ten realizations (black dotted curves), the contour functions (red curve), the lower and upper expectations (vertical broken lines), as well as the lower and upper cdfs (blue curves).

356 The lower and upper cdfs of \tilde{Y} can be computed from (9) and (14) in a similar manner.
 357 The expectation of the logit-normal probability does not have any analytical expression.
 358 Consequently, this is also true for the lower and upper expectations of a logit-normal RFS.
 359 A multidimensional extension of logit-normal RFNs will be studied in Section 3.3.

360 **Example 5.** Figure 5 shows representations of logit-normal RFNs $\tilde{Y}_1 \sim T\tilde{N}(1, 1, 5, \text{logit})$
 361 and $\tilde{Y} \sim T\tilde{N}(-2, 0.1, 2, \text{logit})$, as well as their orthogonal sum $Y_1 \oplus Y_2$.

362 *Parameterized families of transformations.* Whereas the fixed transformations considered
 363 above allow us to define RFNs with various supports, it may be useful to consider more
 364 general transformations belonging to parameterized families. We then obtain a parametrized
 365 family of RFNs with parameters μ , σ^2 , h , and the parameters of the transformation. Such
 366 flexible families may be useful, for instance, in a belief elicitation context where we attempt
 367 to fit an expert's belief statements with a t-GRFN, as will be seen in Section 5.1. The idea
 368 of transforming the normal distribution to obtained parameterized families of distributions
 369 with varying skewness and kurtosis can be traced back, at least, to Ref. [20]. Johnson [20]
 370 actually considers the inverse problem of finding a transformation ψ^{-1} of a random variable
 371 Y such that $X = \psi^{-1}(Y)$ has, approximately, a standard normal distribution. In addition to
 372 the logarithmic and logit transformations, he considers a “system S_U ” based on the following
 373 transformation:

$$X = \psi^{-1}(Y) = \gamma + \delta \sinh^{-1} \frac{Y - \xi}{\lambda}, \quad (18)$$

374 where γ , δ , ξ and λ are four parameters, $\delta > 0$, $\lambda > 0$, and $X \sim N(0, 1)$. Inverting (18), we
 375 get

$$Y = \psi(X) = \xi + \lambda \sinh \frac{X - \gamma}{\delta}. \quad (19)$$

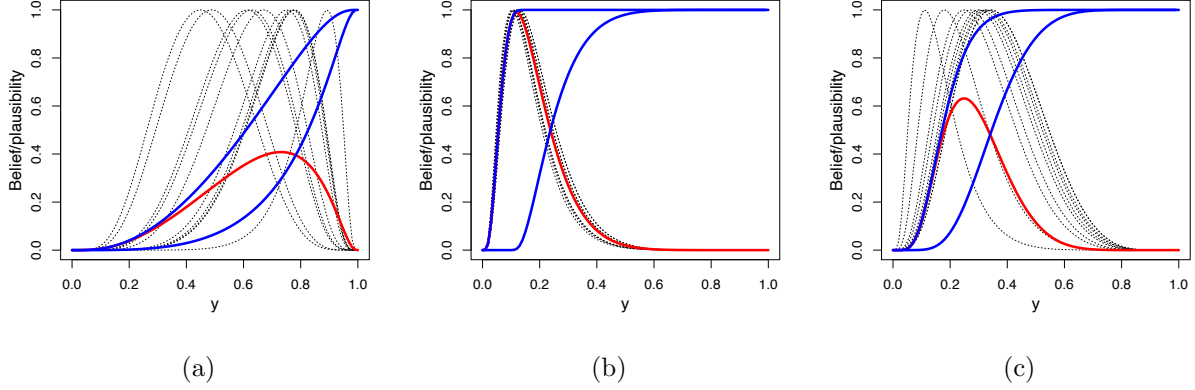


Figure 5: (a) and (b): From left to right, two logit-normal random fuzzy numbers $\tilde{Y} \sim T\tilde{N}(1, 1, 5, \text{logit})$ and $\tilde{Y} \sim T\tilde{N}(-2, 0.1, 2, \text{logit})$; (c): combined lognormal random fuzzy number $\tilde{Y}_1 \oplus \tilde{Y}_2$. For each RFN, we plot ten realizations (black dotted curves), the contour function (red curve), as well as lower and upper cdfs (blue curves).

376 To be consistent with our previous notations, we can rewrite (19) as

$$Y = \psi_{\xi, \lambda}(X) = \xi + \lambda \sinh X, \quad (20)$$

377 where $X \sim N(\mu, \sigma^2)$, and the four parameters are now: μ, σ^2 defining the distribution of X
 378 on the one hand, and ξ and λ defining the transformation on the other hand. This transfor-
 379 mation makes it possible to define a four-parameter family of probability distributions on
 380 the whole real line, with varying skewness and kurtosis. The same transformation applied to
 381 GRFNs defines a parametric family of RFNs with different shapes. We note that Theorem
 382 2 allows us to combine two t-GRFNs $\tilde{Y}_1 \sim T\tilde{N}(\mu_1, \sigma_1^2, h_1, \psi_{\xi, \lambda}^{-1})$ and $\tilde{Y}_2 \sim T\tilde{N}(\mu_2, \sigma_2^2, h_2, \psi_{\xi, \lambda}^{-1})$
 383 with different means, variances and precisions, but the same transformation $\psi_{\xi, \lambda}$.

384 **Example 6.** Figure 6 shows representations of RFNs $\tilde{Y}_1 \sim T\tilde{N}(1, 1, 5, \psi_{\xi, \lambda}^{-1})$ and $\tilde{Y} \sim$
 385 $T\tilde{N}(-2, 0.1, 2, \psi_{\xi, \lambda}^{-1})$ with $\xi = 0$ and $\lambda = 1$, as well as their orthogonal sum $Y_1 \oplus Y_2$. We can
 386 see that the contour functions of \tilde{Y}_1 and \tilde{Y}_2 are, respectively, right-skewed and left-skewed,
 387 while the contour function of $\tilde{Y}_1 \oplus \tilde{Y}_2$ is left-skewed. In general, the positive or negative
 388 skewness depends, for this transformation, on the sign of μ .

389 An even more general approach to create parametric families of probability distributions
 390 was more recently proposed in [2]. Given three random variables X, T and R , Aljarrah et
 391 al. [2] define a new random variable

$$Y = \psi_{RT}(X) = (F_R^{-1} \circ F_T)(X),$$

392 where F_R and F_T are the cdfs of R and T . The cdf of Y is, thus,

$$F_Y(y) = P(Y \leq y) = P(F_T(X) \leq F_R(y)) = P(X \leq F_T^{-1}(F_R(y))) = (F_X \circ F_T^{-1} \circ F_R)(y).$$

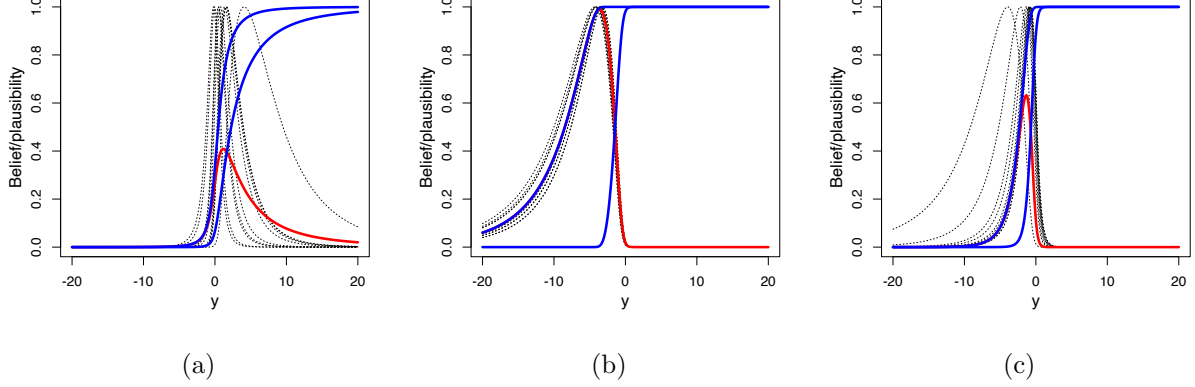


Figure 6: (a) and (b): From left to right, two t-GRFNs $\tilde{Y} \sim T\tilde{N}(1, 1, 5, \psi_{\xi, \lambda}^{-1})$ and $\tilde{Y} \sim T\tilde{N}(-2, 0.1, 2, \psi_{\xi, \lambda}^{-1})$, with $\xi = 0$ and $\lambda = 1$; (c): combined lognormal random fuzzy number $\tilde{Y}_1 \oplus \tilde{Y}_2$. For each RFN, we plot ten realizations (black dotted curves), the contour function (red curve), as well as lower and upper cdfs (blue curves).

393 Taking $X \sim N(\mu, \sigma^2)$ it is possible, using this approach, to define infinitely many parametric
 394 families of probability distributions by choosing different parametric families for T and R .
 395 We note that the support of Y is included in the support of R . Also, choosing $F_T = F_X$ gives
 396 us $Y = R$, while choosing $F_T = F_R$ yields $Y = X$. By extension, we can define parametric
 397 families of RFNs $\tilde{Y} \sim TN(\mu, \sigma^2, h, \psi_{RT}^{-1})$ based on parametric families for R and T .

398 3.3. Transformations of Gaussian Random Vectors

399 The general approach introduced in Section 3.1 can also be applied to GRFVs. Of special
 400 interest is the multivariate extension of the notion of logit-normal RFNs, which is obtained
 401 from a multidimensional normal distribution $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ in \mathbb{R}^{p-1} ($p \geq 2$) and the following
 402 *softmax* transformation from \mathbb{R}^{p-1} to the simplex \mathcal{S}_p of p -dimensional probability vectors:

$$\psi_{\mathcal{S}}(\mathbf{x}) = \left[\frac{\exp(x_1)}{1 + \sum_{j=1}^{p-1} \exp(x_j)}, \dots, \frac{\exp(x_{p-1})}{1 + \sum_{j=1}^{p-1} \exp(x_j)}, \frac{1}{1 + \sum_{j=1}^{p-1} \exp(x_j)} \right]^T, \quad (21)$$

403 with inverse

$$\psi_{\mathcal{S}}^{-1}(\mathbf{y}) = \left[\log\left(\frac{y_1}{y_p}\right), \dots, \log\left(\frac{y_{p-1}}{y_p}\right) \right]^T. \quad (22)$$

404 This transformation is used in [1] to define the multidimensional logistic-normal probability
 405 distribution. Here, given a GRVN $\tilde{X} \sim \tilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{H})$, we can define the random fuzzy vector
 406 $\tilde{Y} = \tilde{\psi}_{\mathcal{S}} \circ \tilde{X}$, where $\tilde{\psi}_{\mathcal{S}}$ is the extension of the softmax transformation (21). We will say that
 407 \tilde{Y} is a *logistic-normal RFV*, and we will write $\tilde{Y} \sim T\tilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{H}, \psi_{\mathcal{S}}^{-1})$. The domain of \tilde{Y} is
 408 the simplex \mathcal{S}_p . Such a RFV can be used to represent beliefs about a vector of probabilities
 409 or proportions.

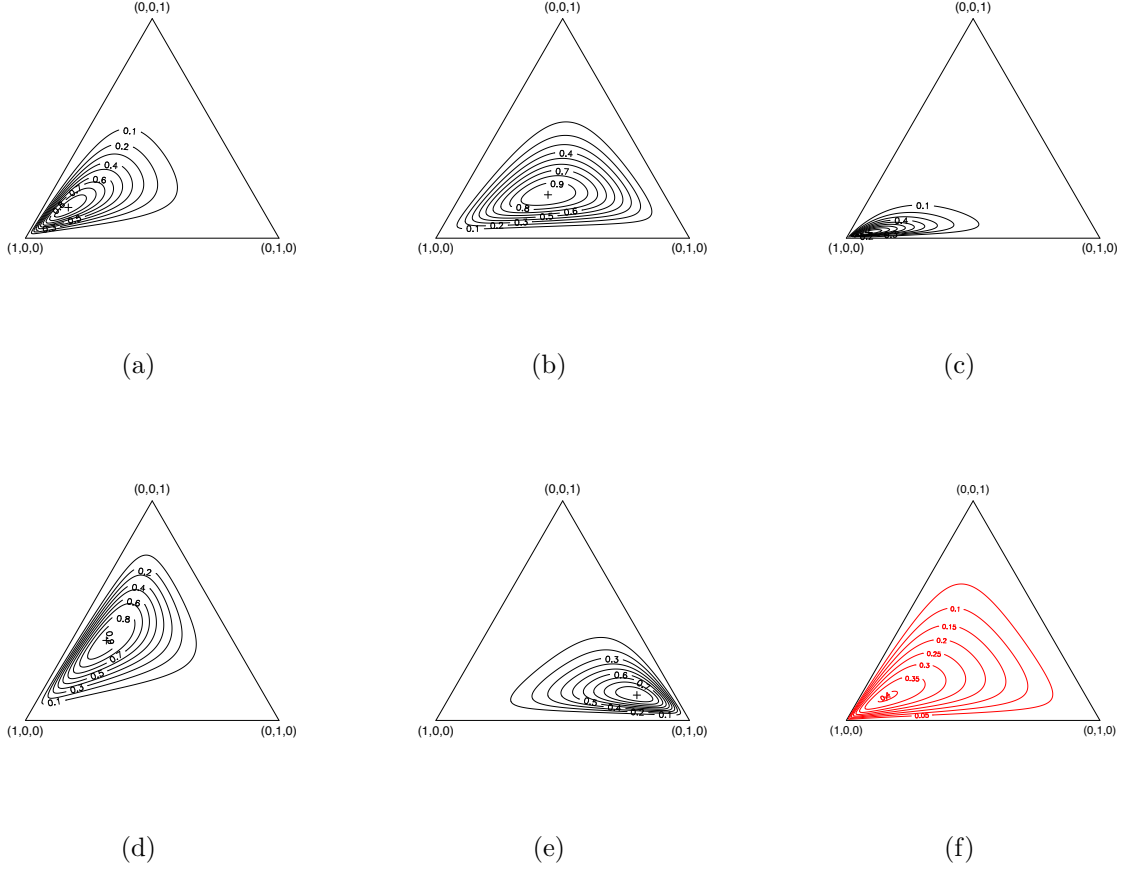


Figure 7: (a)-(e): Five focal fuzzy sets $\tilde{Y}(\omega)$ for a logistic-normal RFV $\tilde{Y} \sim \tilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{H}, \psi_S^{-1})$ (see Example 7); (f): corresponding contour function.

410 **Example 7.** Figures 7a-7e show five focal fuzzy sets $\tilde{Y}(\omega)$ in barycentric coordinates for a
 411 logistic RFV $\tilde{Y} \sim \tilde{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{H}, \psi_S^{-1})$ with $p = 3$, $\boldsymbol{\mu} = (2, 0)^T$,

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1.5 & 0.2 \\ 0.2 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{H} = \begin{pmatrix} 1 & -0.2 \\ -0.2 & 1.5 \end{pmatrix}.$$

412 The corresponding contour function is displayed in Figure 7f.

413 4. Mixtures of Gaussian random fuzzy variables

414 In probability and statistics, finite mixtures of probability distributions and, in particu-
 415 lar, finite mixtures of Gaussians are commonly used to obtain distributions with arbitrarily
 416 complex shapes [23]. In this section, we extend this approach to define mixtures of (trans-
 417 formed) GRFNs and GRFVs. Mixtures of GRFNs and GRFVs will first be defined in

418 Section 4.1. Their properties will then be studied in Section 4.2, and their combination
 419 will be addressed in Section 4.3. A summarization procedure allowing us to approximate a
 420 mixture with a large number of components by a simpler one will be described in Section
 421 4.4. Finally, mixtures of transformed GRFNs and GRFVs will be introduced in Section 4.5.

422 4.1. Definitions

423 *Mixture of GRFNs.* We consider a pair of random variables (M, Z) from a probability space
 424 $(\Omega, \Sigma_\Omega, P)$ to $\mathbb{R} \times \{1, \dots, K\}$, such that the marginal distribution of Z is defined by $P(Z =$
 425 $k) = \pi_k$, $k = 1, \dots, K$, and the conditional distribution of M given $Z = k$ is univariate
 426 normal:

$$M \mid (Z = k) \sim N(\mu_k, \sigma_k^2).$$

427 The marginal distribution of M is, thus, a mixture of K normal distributions. Now, consider
 428 the random fuzzy set $\tilde{X} : \Omega \rightarrow [0, 1]^{\mathbb{R}}$ defined as follows,

$$\tilde{X}(\omega) = \text{GFN} \left(M(\omega), \prod_{k=1}^K h_k^{Z_k(\omega)} \right),$$

429 where $Z_k(\omega) = I(Z(\omega) = k)$, and $I(\cdot)$ is the indicator function. Conditionally on $Z = k$, \tilde{X}
 430 is a GRFN with mean μ_k , variance σ_k^2 and precision h_k :

$$\tilde{X} \mid (Z = k) \sim \tilde{N}(\mu_k, \sigma_k^2, h_k).$$

431 We denote this conditional GRFN by \tilde{X}_k . We say that \tilde{X} is a *mixture GRFN* (m-GRFN)
 432 and we write

$$\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k).$$

433 *Mixture of GRFVs.* Similarly, we can define a mixture of GRFVs by a mapping $\tilde{X} : \Omega \rightarrow$
 434 $[0, 1]^{\mathbb{R}^p}$ with $p \geq 2$ such that

$$\tilde{X}(\omega) = \text{GFV}(\mathbf{M}(\omega), \prod_{k=1}^K \mathbf{H}_k^{Z_k(\omega)}),$$

435 where, as usual, \mathbf{H}_k^0 is the identity matrix, \mathbf{M} is a random vector having a mixture of
 436 multivariate normal distributions,

$$\mathbf{M} \sim \sum_{k=1}^K \pi_k N(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k),$$

437 and \mathbf{H}_k , $k = 1, \dots, K$ are positive definite precision matrices. We will use the following
 438 notation:

$$\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{H}_k).$$

439 *4.2. Properties*

440 *Belief and plausibility functions.* The belief and plausibility functions associated with an
 441 m-GRFN are the weighted sums of, respectively, the belief and plausibility associated with
 442 the components of the mixture. This property is expressed by the following theorem.

Theorem 3. *Let $\mathcal{B}(\mathbb{R})$ be the Borel σ -algebra on \mathbb{R} , and $A \in \mathcal{B}(\mathbb{R})$ be a measurable subset of \mathbb{R} . The degrees of belief and plausibility of A induced by an m-GRFN $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k)$ are*

$$Bel_{\tilde{X}}(A) = \sum_{k=1}^K \pi_k Bel_{\tilde{X}_k}(A) \quad (23a)$$

$$Pl_{\tilde{X}}(A) = \sum_{k=1}^K \pi_k Pl_{\tilde{X}_k}(A), \quad (23b)$$

443 with $\tilde{X}_k \sim \tilde{N}(\mu_k, \sigma_k^2, h_k)$.

Proof. Let us start with (23b). By definition, $Pl_{\tilde{X}}(A)$ is defined as the following expectation,

$$\begin{aligned} Pl_{\tilde{X}}(A) &= \mathbb{E}_{M,Z} \left[\sup_{u \in A} \varphi(u, M, \prod_{k=1}^K h_k^{Z_k}) \right] \\ &= \mathbb{E}_Z \mathbb{E}_{M|Z} \left[\sup_{u \in A} \varphi(u, M, \prod_{k=1}^K h_k^{Z_k}) \right] \\ &= \sum_{k=1}^K \pi_k \mathbb{E}_{M|Z} \left[\sup_{u \in A} \varphi(u, M, h_k) \mid Z = k \right] \\ &= \sum_{k=1}^K \pi_k Pl_{\tilde{X}_k}(A). \end{aligned}$$

444 Eq. (23a) can be proved in the same way, since $Bel_{\tilde{X}}(A)$ is also defined as an expectation. \square

445 Using Theorem 3 and the closed-form expressions given in [14], we can compute the
 446 degrees of belief and plausibility of any real interval $[x, y]$. In particular, setting $x = y$ gives
 447 the following corollary:

448 **Corollary 1.** *The contour function of m-GRFN $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k)$ is*

$$pl_{\tilde{X}}(x) = \sum_{k=1}^K \frac{\pi_k}{\sqrt{1 + h_k \sigma_k^2}} \exp \left(-\frac{h_k (x - \mu_k)^2}{2(1 + h_k \sigma_k^2)} \right). \quad (24)$$

449 *Proof.* Immediate from Theorem 3 and Eq. (8). \square

450 **Example 8.** *Figure 8 shows ten focal sets of an m-GRFN $\tilde{X} \sim 0.4\tilde{N}(-2, 1, 5) + 0.6\tilde{N}(2, 0.1^2, 1)$,
 451 its contour function as well as its lower and upper cdfs.*

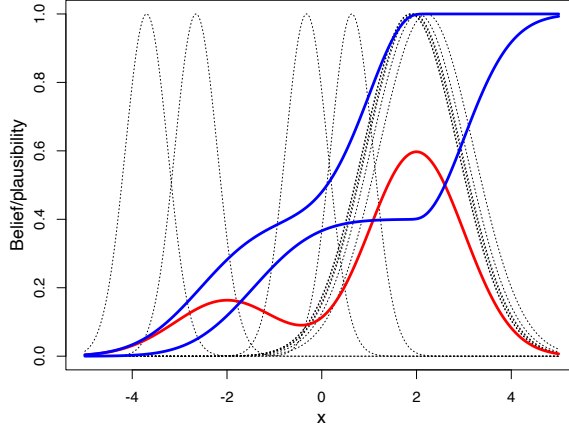


Figure 8: Ten realizations (black dotted curves), contour function (red curve), and lower/upper cdfs (blue curves) of an m-GRFN $\tilde{X} \sim 0.4\tilde{N}(-2, 1, 5) + 0.6\tilde{N}(2, 0.1^2, 1)$.

452 Similar results can be obtained in the same way for mixtures of GRFVs. In particular,
 453 the contour function of a mixture of GRFVs is the weighted sum of the contour functions
 454 of its components, which is expressed by the following proposition.

455 **Proposition 4.** *The contour function of m-GRFV $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{H}_k)$ is*

$$pl_{\tilde{X}}(x) = \sum_{k=1}^K \frac{1}{|\mathbf{I}_p + \boldsymbol{\Sigma}_k \mathbf{H}_k|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T (\mathbf{H}_k^{-1} + \boldsymbol{\Sigma}_k)^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)\right). \quad (25)$$

456 *Lower and upper expectations of an m-GRFN.* The lower and upper expectations of an m-
 457 GRFN can easily be computed from those of its components, as stated in the following
 458 proposition.

Proposition 5. *The lower and upper expectations of m-GRFN $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k)$ are given by*

$$\mathbb{E}_*(\tilde{X}) = \sum_{k=1}^K \pi_k \mu_k - \sum_{k=1}^K \sqrt{\frac{\pi}{2h_k}}, \quad \text{and}$$

$$\mathbb{E}^*(\tilde{X}) = \sum_{k=1}^K \pi_k \mu_k + \sum_{k=1}^K \sqrt{\frac{\pi}{2h_k}}.$$

459 *Proof.* The α -cut of $\tilde{X}(\omega)$ is the closed interval

$${}^\alpha \tilde{X}(\omega) = \left[M(\omega) - \sqrt{\frac{-2 \ln \alpha}{\prod_{k=1}^K h_k^{Z_k(\omega)}}, M(\omega) + \sqrt{\frac{-2 \ln \alpha}{\prod_{k=1}^K h_k^{Z_k(\omega)}}} \right].$$

From (3), the lower expectation of \tilde{X} is

$$\begin{aligned}
\mathbb{E}_*(\tilde{X}) &= \int_0^1 \mathbb{E}_{M,Z} \left[M - \sqrt{\frac{-2 \ln \alpha}{\prod_{k=1}^K h_k^{Z_k}}} \right] d\alpha \\
&= \int_0^1 \mathbb{E}_Z \mathbb{E}_{M|Z} \left[M - \sqrt{\frac{-2 \ln \alpha}{\prod_{k=1}^K h_k^{Z_k}}} \right] d\alpha \\
&= \int_0^1 \sum_{k=1}^K \pi_k \mathbb{E}_{M|Z} \left[M - \sqrt{\frac{-2 \ln \alpha}{h_k}} \mid Z = k \right] d\alpha \\
&= \sum_{k=1}^K \pi_k \int_0^1 \mathbb{E}_{M|Z} \left[M - \sqrt{\frac{-2 \ln \alpha}{h_k}} \mid Z = k \right] d\alpha \\
&= \sum_{k=1}^K \pi_k \left(\mu_k - \sqrt{\frac{\pi}{2h_k}} \right),
\end{aligned}$$

460 where the last equality is derived from (10). The upper expectation of \tilde{X} can be computed
461 in the same way. \square

462 4.3. Combination

463 *Combination of m-GRFNs.* Let us now consider two independent m-GRFNs

$$\tilde{X}_1 \sim \sum_{k=1}^K \pi_{1k} \tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k}) \quad \text{and} \quad \tilde{X}_2 \sim \sum_{\ell=1}^L \pi_{2\ell} \tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell}).$$

464 The following theorem states that their orthogonal sum $\tilde{X}_1 \oplus \tilde{X}_2$ is an m-GRFN.

465 **Theorem 4.** *Given two independent m-GRFNs $\tilde{X}_1 \sim \sum_{k=1}^K \pi_{1k} \tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k})$ and $\tilde{X}_2 \sim$
466 $\sum_{\ell=1}^L \pi_{2\ell} \tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell})$,*

467 1. *The orthogonal sum of \tilde{X}_1 and \tilde{X}_2 is an m-GRFN,*

$$\tilde{X}_1 \oplus \tilde{X}_2 \sim \sum_{k=1}^K \sum_{\ell=1}^L \tilde{\pi}_{k\ell} \left[\tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k}) \oplus \tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell}) \right],$$

468 *with*

$$\tilde{\pi}_{k\ell} = \frac{(1 - \kappa_{k\ell}) \pi_{1k} \pi_{2\ell}}{\sum_{k'\ell'} (1 - \kappa_{k'\ell'}) \pi_{1k'} \pi_{2\ell'}},$$

469 $\kappa_{k\ell}$ denoting the degree of conflict between \tilde{X}_{1k} and $\tilde{X}_{2\ell}$ given by Proposition 1.

470 2. *The degree of conflict between \tilde{X}_1 and \tilde{X}_2 is*

$$\kappa = \sum_{k=1}^K \sum_{\ell=1}^L \kappa_{k\ell} \pi_{1k} \pi_{2\ell}.$$

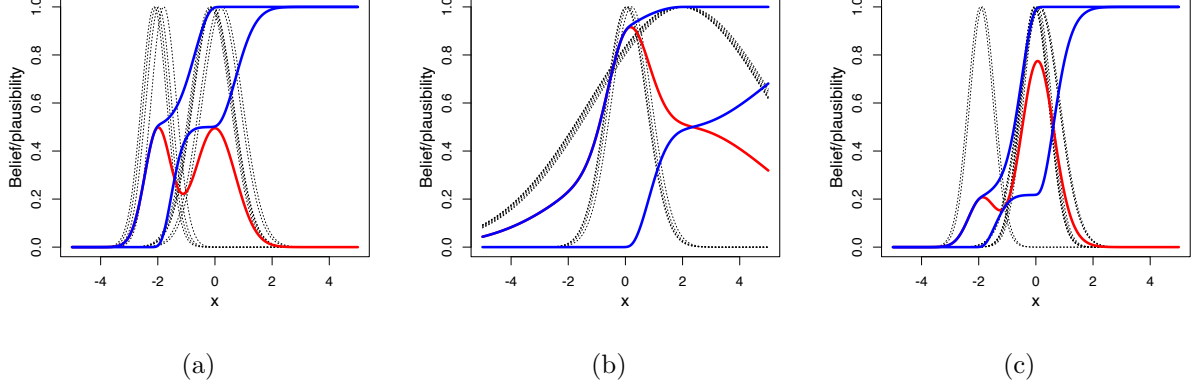


Figure 9: (a) and (b): From left to right, two m-GRFNs $\tilde{X}_1 \sim 0.5\tilde{N}(-2, 0.1^2, 5) + 0.5\tilde{N}(0, 0.1^2, 2)$ and $\tilde{X}_2 \sim 0.5\tilde{N}(0.1, 0.1^2, 2) + 0.5\tilde{N}(2, 0.1^2, 0.1)$; (c): combined mGRFN $\tilde{X}_1 \oplus \tilde{X}_2$. For each RFN, we plot ten realizations (black dotted curves), the contour function (red curve), as well as the lower and upper cdfs (blue curves).

471 *Proof.* See Section Appendix C. □

Example 9. Figure 9 displays two m-GRFNs $\tilde{X}_1 \sim 0.5\tilde{N}(-2, 0.1^2, 5) + 0.5\tilde{N}(0, 0.1^2, 2)$ and $\tilde{X}_2 \sim 0.5\tilde{N}(0.1, 0.1^2, 2) + 0.5\tilde{N}(2, 0.1^2, 0.1)$, as well as their orthogonal sum

$$\begin{aligned} \tilde{X}_1 \oplus \tilde{X}_2 \sim & 0.020\tilde{N}(-1.38, 0.0767^2, 7) + 0.426\tilde{N}(0.05, 0.0707^2, 4) + \\ & 0.197\tilde{N}(-1.92, 0.0980^2, 5.1) + 0.357\tilde{N}(0.0970, 0.0953^2, 2.1). \end{aligned}$$

472 We note that $\tilde{X}_1 \oplus \tilde{X}_2$ has four components but the first one, resulting from the combination of
 473 \tilde{X}_{11} and \tilde{X}_{21} , has a small proportion because of high conflict. When combining a large number
 474 of m-GRFNs, the number of components grows exponentially. However, the combined m-
 475 GRFN can be approximated by a simpler one, using a technique that will be introduced in
 476 Section 4.4.

477 *Combination of m-GRFVs.* Theorem 4 can easily be generalized to GRFVs, using the re-
 478 sults in presented in [14, Proposition 13]. Basically, when combining two GRFVs, each
 479 component of the first GRFV is combined with each component of the second GRFV, and
 480 the proportions are adjusted based on degrees of conflict. This result is formally stated in
 481 the next theorem.

482 **Theorem 5.** Let $\tilde{X}_1 \sim \sum_{k=1}^K \pi_{1k} \tilde{N}(\boldsymbol{\mu}_{1k}, \boldsymbol{\Sigma}_{1k}^2, \mathbf{H}_{1k})$, and $\tilde{X}_2 \sim \sum_{\ell=1}^L \pi_{2\ell} \tilde{N}(\boldsymbol{\mu}_{2\ell}, \boldsymbol{\Sigma}_{2\ell}^2, \mathbf{H}_{2\ell})$ be
 483 two p -dimensional m-GRFVs and assume that the precision matrices \mathbf{H}_{1k} and $\mathbf{H}_{2\ell}$ are all
 484 positive definite. Then:

485 1. The orthogonal sum of \tilde{X}_1 and \tilde{X}_2 is an m -GRFV,

$$\tilde{X}_1 \oplus \tilde{X}_2 \sim \sum_{k=1}^K \sum_{\ell=1}^L \tilde{\pi}_{k\ell} \left[\tilde{N}(\boldsymbol{\mu}_{1k}, \boldsymbol{\Sigma}_{1k}^2, \mathbf{H}_{1k}) \oplus \tilde{N}(\boldsymbol{\mu}_{2\ell}, \boldsymbol{\Sigma}_{2\ell}^2, \mathbf{H}_{2\ell}) \right],$$

486 with

$$\tilde{\pi}_{k\ell} = \frac{(1 - \kappa_{k\ell})\pi_{1k}\pi_{2\ell}}{\sum_{k'\ell'} (1 - \kappa_{k'\ell'})\pi_{1k'}\pi_{2\ell'}},$$

487 where $\kappa_{k\ell}$ is the degree of conflict between \tilde{X}_{1k} and $\tilde{X}_{2\ell}$ given by (A.1).

488 2. The degree of conflict between \tilde{X}_1 and \tilde{X}_2 is

$$\kappa = \sum_{k=1}^K \sum_{\ell=1}^L \kappa_{k\ell}\pi_{1k}\pi_{2\ell}.$$

489 *Proof.* The theorem can be proved by following a similar line of reasoning as in the proofs
490 of Lemma 3 and Theorem 4, using the results in Lemma 2 and Proposition 13 in [14]. \square

491 4.4. Summarization of an m -GRFN

492 As already mentioned, the number of components grows exponentially when combining
493 m -GRFNs: if all of N m -GRFNs have the same number K of components, the combined
494 m -GRFN has, in general, K^N components. However, components of a combined m -GRFN
495 resulting from the combination of highly conflicting GRFNs have a very small proportion.
496 When combining many m -GRFNs, we can expect to obtain a large number of components,
497 many of which will have a proportion close to zero. A similar problem occurs when combining
498 a large number of mass functions, in which case we often observe a proliferation of focal sets
499 with very small masses. Several strategies have been proposed for controlling the number
500 of focal sets by transferring some masses to supersets, resulting in an outer approximation
501 of the original mass function [22, 6, 19]. The simplest approach is to transfer all masses
502 less than some threshold to the union of the corresponding focal sets [22]. Here, we do not
503 have a notion of “union” for GRFNs that would result in a GRFN, but we can transfer
504 the small proportions to a vacuous GRFN $\tilde{X}_0 \sim \tilde{N}(\mu, \sigma^2, 0)$ with arbitrary μ and σ^2 . This
505 procedure, formally described in Algorithm 1, yields a conservative approximation of the
506 original m -GRFN. It is illustrated by the following example.

507 **Example 10.** Consider the following three m -GRFNs, each one with three components:

$$\tilde{X}_1 \sim \frac{1}{3}\tilde{N}(-2, (0.1)^2, 2) + \frac{1}{3}\tilde{N}(0, (0.1)^2, 2) + \frac{1}{3}\tilde{N}(2, (0.1)^2, 2),$$

$$\tilde{X}_2 \sim \frac{1}{3}\tilde{N}(-2.1, (0.5)^2, 2) + \frac{1}{3}\tilde{N}(0.1, (0.5)^2, 2) + \frac{1}{3}\tilde{N}(1/9, (0.5)^2, 2),$$

$$\tilde{X}_3 \sim \frac{1}{3}\tilde{N}(-1.9, (0.5)^2, 5) + \frac{1}{3}\tilde{N}(-0.1, (0.5)^2, 5) + \frac{1}{3}\tilde{N}(2.1, (0.5)^2, 5).$$

Algorithm 1 Summarization of a m-GRFN.

Input: m-GRFN $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k)$, threshold $\epsilon < 1$
 Reorder the components of \tilde{X} such that $\pi_{(1)} \leq \pi_{(2)} \leq \dots \leq \pi_{(K)}$
if $\pi_{(1)} \geq \epsilon$ **then**
 $\tilde{X}' \leftarrow \tilde{X}$
else
 Find the largest k_0 such that $\sum_{k=1}^{k_0} \pi_{(k)} \leq \epsilon$
 $\pi_0 \leftarrow \sum_{k=1}^{k_0} \pi_{(k)}$
 $\tilde{X}' \leftarrow \pi_0 \tilde{N}(0, 1, 0) + \sum_{k=k_0+1}^K \pi_{(k)} \tilde{N}(\mu_{(k)}, \sigma_{(k)}^2, h_{(k)})$
end if
Output: Approximate m-GRFN \tilde{X}'

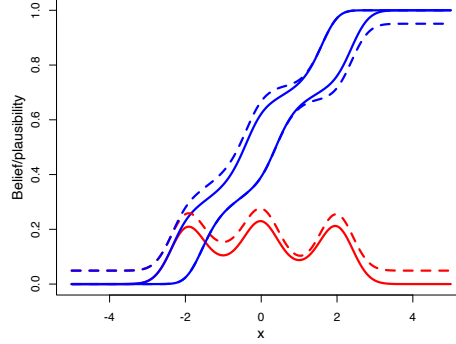


Figure 10: Contour functions (red curves), and lower/upper cdfs (blue curves) of m-GRFN \tilde{X} with 27 components in Example 10 (solid lines) and its approximation \tilde{X}' with only 13 components (broken lines).

510 *The combined m-GRFN $\tilde{X} = \tilde{X}_1 \oplus \tilde{X}_2 \oplus \tilde{X}_3$ has 27 components. Summarizing \tilde{X} with a*
 511 *threshold $\epsilon = 0.05$ yields an n-GRFN with 13 components. The contour function as well as*
 512 *the lower and upper cdfs of \tilde{X} and \tilde{X}' are shown in Figure 10. As we can see, \tilde{X}' is a good*
 513 *conservative approximation of \tilde{X} , while having less than half as many components.*

514 4.5. Mixtures of transformed GRFNs

515 *Definition.* The ideas developed in Section 3 and in the current section can naturally be
 516 combined to define mixtures of transformed fuzzy numbers. Let us consider a one-to-one
 517 mapping ψ from \mathbb{R} to $\Lambda \subseteq \mathbb{R}$ and, as in Section 4.1, an m-GRFN $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\mu_k, \sigma_k^2, h_k)$,
 518 originating from a pair of random variables (M, Z) from a probability space $(\Omega, \Sigma_\Omega, P)$ to
 519 $\mathbb{R} \times \{1, \dots, K\}$, such that $P(Z = k) = \pi_k$, $k = 1, \dots, K$, and $M \mid (Z = k) \sim N(\mu_k, \sigma_k^2)$,
 520 $k = 1, \dots, K$. From Proposition 2, the composed mapping

$$\tilde{Y}(\omega) = (\tilde{\psi} \circ \tilde{X})(\omega) = \tilde{\psi} \left[\text{GFN}(M(\omega), \prod_{k=1}^K h_k^{Z_k(\omega)}) \right],$$

521 where $Z_k(\omega) = I(Z(\omega) = k)$, is a RFS. Conditionally on $Z = k$, \tilde{Y} is a t-GRFN with mean
 522 μ_k , variance σ_k^2 and precision h_k :

$$\tilde{Y} \mid (Z = k) \sim T\tilde{N}(\mu_k, \sigma_k^2, h_k, \psi^{-1}).$$

523 We say that \tilde{Y} is a mixture of transformed GRFNs (mt-GRFN) and we write

$$\tilde{Y} \sim \sum_{k=1}^K \pi_k T\tilde{N}(\mu_k, \sigma_k^2, h_k, \psi^{-1}).$$

524 In a similar way, we can define a mixture of transformed GRFVs (mt-GRFV) as the trans-
 525 formation of an m-GRFV $\tilde{X} \sim \sum_{k=1}^K \pi_k \tilde{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{H}_k)$ by a one-to-one mapping ψ from \mathbb{R}^p
 526 to $\Lambda \subseteq \mathbb{R}^p$ and write

$$\tilde{Y} = \tilde{\psi} \circ \tilde{X} \sim \sum_{k=1}^K \pi_k T\tilde{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{H}_k, \psi^{-1}).$$

527 *Properties.* The properties of mt-GRFNs can be derived directly from those of m-GRFNs
 528 and t-GRFNs. Some of them are described in the following propositions. Similar properties
 529 hold for mt-GRFVs.

530 **Proposition 6.** *Let $\tilde{Y} \sim \sum_{k=1}^K \pi_k T\tilde{N}(\mu_k, \sigma_k^2, h_k, \psi^{-1})$ be an mt-GRFN, where ψ is a one-
 531 to-one mapping from \mathbb{R} to $\Lambda \subseteq \mathbb{R}$. For any event $A \in \Sigma_\Lambda$,*

$$Bel(A) = \sum_{k=1}^K \pi_k Bel_{\tilde{X}_k}(\psi^{-1}(A)), \quad \text{and} \quad Pl(A) = \sum_{k=1}^K \pi_k Pl_{\tilde{X}_k}(\psi^{-1}(A)),$$

532 where $\tilde{X}_k \sim \tilde{N}(\mu_k, \sigma_k^2, h_k)$. In particular the contour function of \tilde{Y} is

$$pl_{\tilde{Y}}(y) = \sum_{k=1}^K \frac{\pi_k}{\sqrt{1 + h_k \sigma_k^2}} \exp\left(-\frac{h_k(\psi^{-1}(y) - \mu_k)^2}{2(1 + h_k \sigma_k^2)}\right).$$

533 *Proof.* Immediate from Theorem 1, Theorem 3 and Corollary 1. □

534 **Proposition 7.** *Let $\tilde{Y}_1 \sim \sum_{k=1}^K \pi_{1k} T\tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k}, \psi^{-1})$ and $\tilde{Y}_2 \sim \sum_{\ell=1}^L \pi_{2\ell} T\tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell}, \psi^{-1})$
 535 be two mt-GRFNs, where ψ is a one-to-one mapping from \mathbb{R} to $\Lambda \subseteq \mathbb{R}$. Their orthogonal
 536 sum is the mt-GRFN*

$$\tilde{Y}_1 \oplus \tilde{Y}_2 \sim \sum_{k=1}^K \sum_{\ell=1}^L \tilde{\pi}_{k\ell} T\tilde{N}(\tilde{\mu}_{k\ell}, \tilde{\sigma}_{k\ell}^2, h_{1k} + h_{2\ell}, \psi^{-1}),$$

537 where $\tilde{\mu}_{k\ell}$ and $\tilde{\sigma}_{k\ell}^2$ are the mean and variance of the orthogonal sum $\tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k}) \oplus$
 538 $\tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell})$ given by (11), and

$$\tilde{\pi}_{k\ell} = \frac{(1 - \kappa_{k\ell})\pi_{1k}\pi_{2\ell}}{\sum_{k'\ell'} (1 - \kappa_{k'\ell'})\pi_{1k'}\pi_{2\ell'}},$$

539 where $\kappa_{k\ell}$ is the degree of conflict between $\tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k})$ and $\tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell})$ given by Propo-
 540 sition 1.

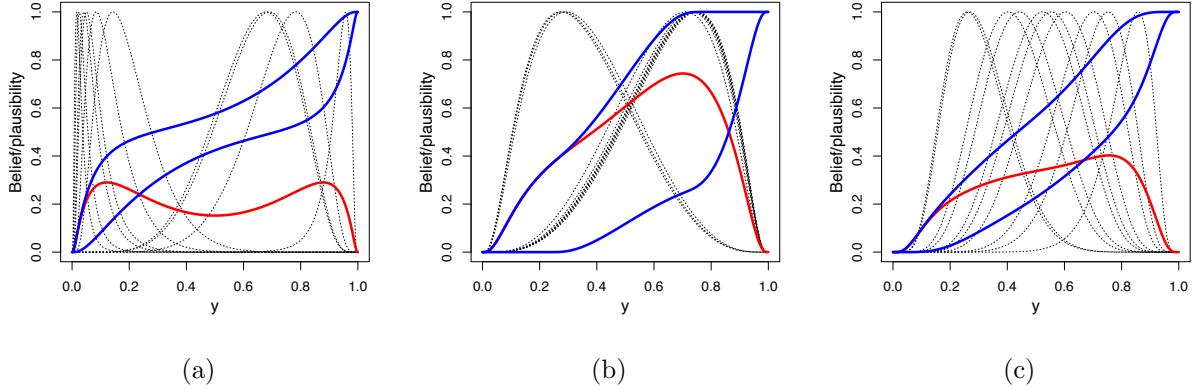


Figure 11: (a) and (b): From left to right, two mixtures of logit-normal RFNs $\tilde{Y}_1 \sim 0.5T\tilde{N}(2, 1, 2, \text{logit}) + 0.5T\tilde{N}(-2, 1, 2, \text{logit})$ and $\tilde{Y}_2 \sim 0.3T\tilde{N}(-1, 0.1^2, 1, \text{logit}) + 0.7T\tilde{N}(1, 0.1^2, 1, \text{logit})$; (c): orthogonal sum $\tilde{Y}_1 \oplus \tilde{Y}_2$. For each RFN, we plot ten realizations (black dotted curves), the contour function (red curve), as well as the lower and upper cdfs (blue curves).

541 *Proof.* Immediate from Theorems 2 and 4. □

542 **Example 11.** Figure 11 displays two mixtures of logit-normal RFNs $\tilde{Y}_1 \sim 0.5T\tilde{N}(2, 1, 2, \text{logit}) +$
543 $0.5T\tilde{N}(-2, 1, 2, \text{logit})$ and $\tilde{Y}_2 \sim 0.3T\tilde{N}(-1, 0.1^2, 1, \text{logit}) + 0.7T\tilde{N}(1, 0.1^2, 1, \text{logit})$, as well as
544 their orthogonal sum $\tilde{Y}_1 \oplus \tilde{Y}_2$.

545 5. Applications

546 In this section, we will discuss two applications of the models introduced in this paper.
547 Belief elicitation will first be addressed in Section 5.1. The use of m-GRFNs to represent
548 weak prior knowledge in generalized Bayesian inference will then be discussed in Section 5.2.

549 5.1. Elicitation

550 According to Garthwaite et al. [17], “elicitation is the process of formulating a person’s
551 knowledge and beliefs about one or more uncertain quantities into a (joint) probability dis-
552 tribution for those quantities”. There is, however, no reason to limit oneself to probability
553 distributions for representing a person’s beliefs. In this section, we briefly discuss the ap-
554 plication of the RFN models introduced in this paper to represent expert beliefs about a
555 numerical quantity. It is clear that an in-depth treatment of this topic would require to
556 delve into complex methodological issues arising when interviewing human experts. Here,
557 we will only be concerned with the use of already obtained plausibility assessments to fit
558 the parameters of a RFN.

559 For single-expert probabilistic elicitation, the two main approaches are the *fixed interval*
560 *method*, in which the expert is asked to give his subjective probabilities for some fixed
561 intervals, and the *variable interval method* in which the expert is invited to provide points

562 corresponding to specified percentiles of his subjective probability distribution. Assuming
 563 the latter approach is used, a sequence of questions for eliciting an expert's beliefs about a
 564 numerical quantity X could be the following:

- 565 1. What is the most plausible value m_0 of X ?
- 566 2. Given two numbers $0 < \alpha < \beta < 1$ (such as, e.g., $\alpha = 0.1$ and $\beta = 0.5$), give values x_α
 567 and x_β such that $Pl(X \leq x_\alpha) = \alpha$ and $Pl(X \leq x_\beta) = \beta$.
- 568 3. Give values x'_α and x'_β such that $Pl(X > x'_\alpha) = \alpha$ and $Pl(X > x'_\beta) = \beta$.

569 This procedure yields a maximum-plausibility value m_0 and the plausibility degrees of four
 570 intervals $(-\infty, x_\alpha]$, $(-\infty, x_\beta]$, $[x'_\alpha, +\infty)$ and $[x'_\beta, +\infty)$. Whatever the details of the elicitation
 571 procedure, we can assume that we have obtained m_0 and the plausibilities pl_1, \dots, pl_n of n
 572 real intervals I_1, \dots, I_n .

573 Several parametric families of RFN proposed in Sections 3.3 and 4 could be fitted to
 574 such data. As the number n of intervals will typically be small, simpler models should be
 575 preferred. Mixture models may not be the most suitable because they depend on many
 576 parameters (at least seven for the simplest two-component case) and they can yield mul-
 577 timodal contour functions. Let us, thus, consider a parameterized family of t-GRFNs
 578 $\tilde{X}_\theta \sim T\tilde{N}(\mu, \sigma^2, h, \psi_\eta^{-1})$, where η is a vector of parameters for the transformation func-
 579 tion ψ , and $\theta = (\mu, \sigma^2, h, \eta)$ is the vector of all parameters. We can then identify θ by
 580 minimizing the following mean squared error function:

$$\text{MSE}(\theta) = \sum_{i=1}^n (Pl_{\tilde{X}_\theta}(I_i) - pl_i)^2. \quad (26)$$

581 subject to the constraint $\psi_\eta(\mu) = m_0$. The following example illustrate this approach with
 582 the family of t-GRFNs defined by transformation $\psi_\eta = \psi_{\xi, \lambda}$ given by (20).

583 **Example 12.** Assume that an expert gives us $m_0 = 1$ as the most plausible value of X , and
 584 the following plausibility assessments:

$$Pl(X \leq -7) = 0.1, \quad Pl(X \leq -1) = 0.5, \quad Pl(X > 2) = 0.5, \quad Pl(X > 5) = 0.1.$$

585 The constraint $\psi_{\xi, \lambda}(\mu) = m_0$ gives us $\mu = \sinh^{-1}(\frac{m_0 - \xi}{\lambda})$. Substituting μ by its expression
 586 as a function of ξ and λ and minimizing (26) w.r.t. σ^2 , h , ξ and λ yields the following
 587 estimates:

$$\hat{\sigma}^2 = 0.73, \quad \hat{h} = 10.90, \quad \hat{\xi} = 3.37, \quad \hat{\lambda} = 2.71,$$

588 and $\hat{\mu} = -0.79$. Figure 12 shows of the lower and upper cdf of the fitted t-GRFN, as well
 589 as its contour function. By comparison, we also show the corresponding functions for a
 590 GRFN fitted on the same data. Obviously, a GRFN is a poor fit given the asymmetry of the
 591 plausibility assessments.

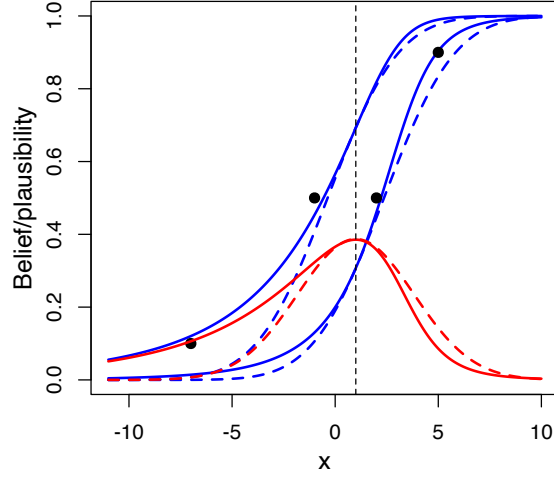


Figure 12: Contour functions (red curves), and lower/upper cdfs (blue curves) of the fitted t-GRFN $\tilde{X}_{\hat{\theta}} \sim T\tilde{N}(\hat{\mu}, \hat{\sigma}^2, \hat{h}, \psi_{\hat{\xi}, \hat{\lambda}}^{-1})$ of Example 12 (solid lines) and a fitted GRFN (broken lines). The data points are shown as black dots, and the vertical broken lines marks the most plausible value m_0 .

5.2. Generalized Bayesian inference

Let us consider a statistical model in which observed data X are drawn randomly from a probability distribution P_{θ} , where $\theta \in \Theta$ is some unknown parameter. Having observed a realization x of X , we define the likelihood function as the mapping

$$L(\cdot; x) : \Theta \rightarrow [0, +\infty)$$

$$\theta \mapsto L(\theta; x) = f(x; \theta),$$

where $f(x; \theta)$ denotes the probability mass or density function of X . Assuming $\sup_{\theta \in \Theta} L(\theta; x) < +\infty$, the relative likelihood function can be defined as

$$\tilde{L}_x : \Theta \rightarrow [0, +\infty)$$

$$\theta \mapsto \tilde{L}_x(\theta) = \frac{L(\theta; x)}{\sup_{\theta' \in \Theta} L(\theta'; x)}.$$

In [8], we showed that \tilde{L}_x can be interpreted as the membership function of a fuzzy subset of Θ (the fuzzy subset of “likely” values of θ after observing x). Equivalently, \tilde{L}_x can be seen as a constant RFS. As shown in [8], \tilde{L}_x is, in some sense, the least committed RFS verifying the following two requirements:

1. Compatibility with Bayesian inference: let P_0 be a prior probability measure on Θ ; combining it with \tilde{L}_x using the product-intersection rule yields the Bayesian posterior distribution $P(\cdot|x)$:

$$P_0 \oplus \tilde{L}_x = P(\cdot|x). \quad (27)$$

600 2. Combination of independent observations: if x and y are realizations of two independent
601 observations X and Y , then the RFS induced by (x, y) is the orthogonal sum of
602 the RFS induced by x , and that induced by y :

$$\tilde{L}_x \oplus \tilde{L}_y = \tilde{L}_{x,y}.$$

603 From the perspective of a general theory of epistemic random fuzzy sets, there is no reason
604 to limit oneself to probability distributions for representing prior information. Indeed, prior
605 knowledge is often vague and the relevance of representing it by precise probabilities is
606 questionable. Equation (27) can, thus, be generalized as

$$\tilde{\theta}_0 \oplus \tilde{L}_x = \tilde{\theta}_x, \quad (28)$$

607 where $\tilde{\theta}_0$ is a RFS representing *weak prior information*, and $\tilde{\theta}_x$ is posterior RFS resulting from
608 the combination of prior information with observations. Considering the case where $\Theta \subseteq \mathbb{R}$,
609 (28) lends itself to easy computation if \tilde{L}_x is a GFN: in this case, modeling prior information
610 as a GRFN $\tilde{\theta}_0 \sim \tilde{N}(\mu_0, \sigma_0^2, h_0)$ or, more generally, an m-GRFN $\tilde{\theta}_0 \sim \sum_{k=1}^K \pi_{0k} \tilde{N}(\mu_{0k}, \sigma_{0k}^2, h_{0k})$
611 will result, respectively, in a posterior GRFN $\tilde{\theta}_x \sim \tilde{N}(\mu_x, \sigma_x^2, h_x)$ or a posterior m-GRFN
612 $\tilde{\theta}_x \sim \sum_{k=1}^K \pi_{xk} \tilde{N}(\mu_{xk}, \sigma_{xk}^2, h_{xk})$.

Except in a few simple cases, the relative likelihood function is not exactly a GFN,
but this model can be used as an approximation [7]. Indeed, a Taylor series expansion of
 $\log \tilde{L}_x(\theta)$ about the maximum likelihood estimate (MLE) $\hat{\theta}$ up to the second order gives us
[28, p. 33]:

$$\log \tilde{L}_x(\theta) = \log \tilde{L}_x(\hat{\theta}) + (\theta - \hat{\theta}) \left. \frac{\partial \log \tilde{L}_x(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} + \frac{1}{2} (\theta - \hat{\theta})^2 \left. \frac{\partial^2 \log \tilde{L}_x(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} + \dots$$

613 The first term on the right-hand is equal to zero by definition, and the second term is zero
614 in the usual case where $\hat{\theta}$ is a stationary point of the likelihood function. Neglecting the
615 higher-order terms, we get the approximation

$$\tilde{L}_x(\theta) \approx \exp \left[-\frac{1}{2} I(\hat{\theta}; x) (\theta - \hat{\theta})^2 \right], \quad (29)$$

616 i.e., $\tilde{L}_x \approx \text{GFN}(\hat{\theta}, I(\hat{\theta}; x))$, where the precision $I(\hat{\theta}; x)$ is the observed Fisher information

$$I(\hat{\theta}; x) = - \left. \frac{\partial^2 \log \tilde{L}_x(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} = - \left. \frac{\partial^2 \log L(\theta; x)}{\partial \theta^2} \right|_{\theta=\hat{\theta}}.$$

617 In the multidimensional case where $\Theta \subseteq \mathbb{R}^p$, the same line of reasoning as above yields
618 $\tilde{L}_x \approx \text{GFV}(\hat{\theta}, I(\hat{\theta}; x))$, where $I(\hat{\theta}; x)$ is the observed information matrix.

As noted in [28], the quality of the normality assumption can sometimes be improved for
small samples by applying some transformation to the parameter. Let $\theta = \psi(\delta)$ for some one-
to-one differentiable mapping ψ and alternative parameter δ , and assume that the relative

likelihood as a function of δ , $\tilde{L}_x^{(\delta)}(\delta)$, is approximately Gaussian, $\tilde{L}_x^{(\delta)}(\delta) \approx \text{GFN}(\hat{\delta}, I(\hat{\delta}, x))$. Then, the likelihood as a function of θ is $\tilde{L}_x^{(\theta)}(\theta) = \tilde{\psi}(\tilde{L}_x^{(\delta)})(\theta) = \tilde{L}_x^{(\delta)}(\psi^{-1}(\theta))$, and it can be approximated by

$$\tilde{L}_x^{(\delta)}(\psi^{-1}(\theta)) \approx \exp \left[-\frac{1}{2} I(\hat{\delta}; x) \left(\psi^{-1}(\theta) - \psi^{-1}(\hat{\theta}) \right)^2 \right] \quad (30a)$$

$$\approx \exp \left[-\frac{1}{2} I(\hat{\theta}; x) [\psi'(\hat{\delta})]^2 \left(\psi^{-1}(\theta) - \psi^{-1}(\hat{\theta}) \right)^2 \right], \quad (30b)$$

619 where we have used the equality $I(\hat{\delta}; x) = I(\hat{\theta}; x) [\psi'(\hat{\delta})]^2$ (see [28, p. 35]), and ψ' denotes
 620 the first derivative of ψ . The fuzzy set $\tilde{L}_x^{(\theta)}$ can be combined with a t-GRFN prior $\tilde{\theta}_0 \sim$
 621 $T\tilde{N}(\mu_0, \sigma_0^2, h_0, \psi^{-1})$ or, more generally, an mt-GRFN $\tilde{\theta}_0 \sim \sum_{k=1}^K \pi_{0k} T\tilde{N}(\mu_{0k}, \sigma_{0k}^2, h_{0k}, \psi^{-1})$ to
 622 obtain, respectively, a posterior t-GRFN $\tilde{\theta}_x \sim T\tilde{N}(\mu_x, \sigma_x^2, h_x, \psi^{-1})$, or a posterior mt-GRFN

$$\tilde{\theta}_x \sim \sum_{k=1}^K \pi_{xk} T\tilde{N}(\mu_{xk}, \sigma_{xk}^2, h_{xk}, \psi^{-1}).$$

623 **Example 13.** Consider an iid sample $x = (x_1, \dots, x_n)$ from a Poisson distribution with
 624 mean θ . Let $t = \sum_{i=1}^n x_i$. The likelihood function is

$$L(\theta; x) = \prod_{i=1}^n \frac{\theta^{x_i} \exp(-\theta)}{x_i!} = \frac{\theta^t \exp(-n\theta)}{\prod x_i!},$$

625 and the MLE of θ is $\hat{\theta} = t/n$, which gives us the following expression for the relative likeli-
 626 hood:

$$\tilde{L}_x(\theta) = \left(\frac{\theta}{\hat{\theta}} \right)^t \exp \left[n(\hat{\theta} - \theta) \right].$$

627 It can easily be shown that $I(\hat{\theta}; x) = n/\hat{\theta}$, so that (29) yields

$$\tilde{L}_x(\theta) \approx \exp \left[-\frac{n}{2\hat{\theta}} (\theta - \hat{\theta})^2 \right] = \exp \left[-\frac{n}{2} \left(\frac{\theta}{\hat{\theta}} - 1 \right)^2 \right]. \quad (31)$$

628 Alternatively, let $\theta = \exp(\delta)$. Approximation (30) gives us

$$\tilde{L}_x(\theta) \approx \exp \left[-\frac{n}{2\hat{\theta}} \exp(2\hat{\theta}) (\log \theta - \log \hat{\theta})^2 \right]. \quad (32)$$

629 Figure 13 displays the exact relative likelihood $\tilde{L}_x(\theta)$ for $n = 10$ and $t = 30$, as well as its
 630 approximations (31) and (32). As we can see, the lognormal approximation (32) is more
 631 accurate, and it takes into account the positivity of θ .

632 Figure 14a shows two priors: a lognormal Bayesian prior $\tilde{\theta}_0 \sim T\tilde{N}(0.5, 0.1, +\infty, \log)$,
 633 represented by its cdf (blue dotted line) and a weaker lognormal t-GRFN prior

$$\tilde{\theta}'_0 \sim T\tilde{N}(0.5, 0.1, 10, \log),$$

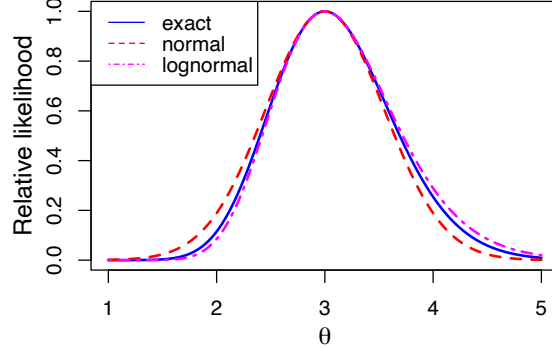


Figure 13: Relative likelihood $\tilde{L}_x(\theta)$ and its normal and lognormal approximations for an iid sample for the Poisson distribution with mean θ , for $n = 10$ observations and $t = 30$.

634 represented by its contour function (red solid line), its upper and lower cdfs (blues solid
635 lines) and 10 realizations (black dotted lines). We also show the lognormal approximation
636 to the relative likelihood as a red broken line. Figure 14b shows the corresponding lognormal
637 t -GRFN posteriors $\tilde{\theta}_x = \tilde{\theta}_0 \oplus \tilde{L}_x$ (equal to the posterior probability distribution) and $\tilde{\theta}'_x =$
638 $\tilde{\theta}'_0 \oplus \tilde{L}_x$, as well as the approximated relative likelihood. As expected, the posterior lognormal
639 t -GRFN corresponding to the weaker prior is more imprecise, and closer to the relative
640 likelihood. It is clear from Figure 14b that our approach is not a robust Bayes method:
641 although the Bayes prior cdf is comprised between the lower and upper weak prior cdfs, this
642 is not the case for the posteriors. In the extreme situation where the prior is vacuous,
643 $\tilde{\theta}_0^v \sim T\tilde{N}(0, 1, 0, \log)$, the posterior is the relative likelihood, as $\tilde{\theta}_0^v \oplus \tilde{L}_x = \tilde{L}_x$. This is in
644 contrast with the robust Bayes approach, which yields a (useless) vacuous posterior with the
645 same vacuous prior information.

646 To illustrate the possibility of taking into account more complex prior information, we
647 show a bimodal mt -GRFN prior

$$\tilde{\theta}_0 \sim 0.5 T\tilde{N}(0.5, 0.1, 30, \log) + 0.5 T\tilde{N}(1.5, 0.01, 20, \log)$$

648 in Figure 15a, and the corresponding posterior mt -GRFN

$$\tilde{\theta}_x = \tilde{\theta}_0 \oplus \tilde{L}_x = 0.29 T\tilde{N}(0.84, 0.047, 60, \log) + 0.71 T\tilde{N}(1.26, 0.0040, 50, \log).$$

649 in Figure 15b. In practice, the prior RFS could be elicited by a method similar to that
650 proposed in Section 5.1.

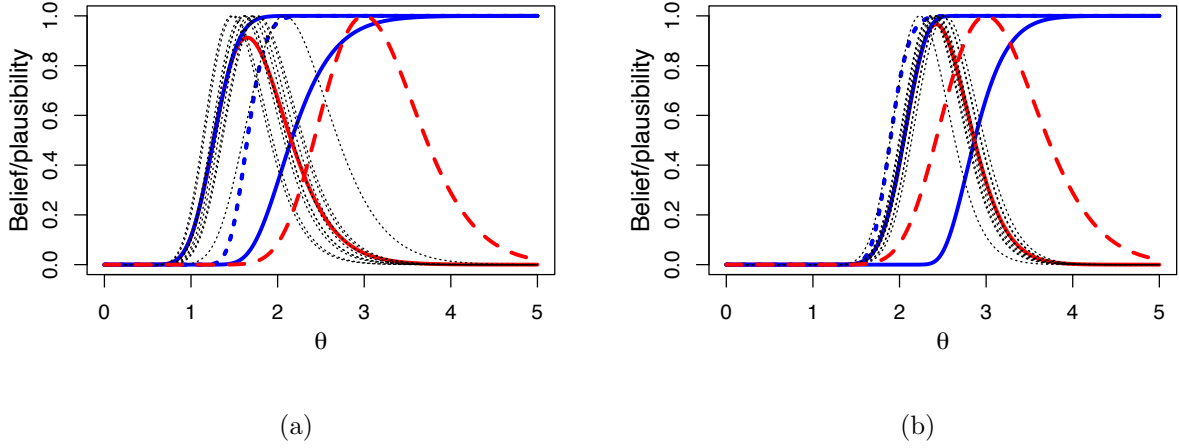


Figure 14: (a) Lognormal Bayesian prior $\tilde{\theta}_0 \sim T\tilde{N}(0.5, 0.1, +\infty, \log)$ (blue dotted line), lognormal t-GRFN prior $\tilde{\theta}'_0 \sim T\tilde{N}(0.5, 0.1, 10, \log)$ (contour function: red solid line, upper and lower cdfs: blues solid lines, and 10 realizations: black dotted lines), and lognormal approximation to the relative likelihood \tilde{L}_x (red broken line); (b): Bayesian lognormal posterior distribution $\tilde{\theta}_x = \tilde{\theta}_0 \oplus \tilde{L}_x$ (blue dotted line), lognormal t-GRFN posterior $\tilde{\theta}'_x = \tilde{\theta}'_0 \oplus \tilde{L}_x$ (contour function: red solid line, upper and lower cdfs: blues solid lines, and 10 realizations: black dotted lines), and lognormal approximation to the relative likelihood (red broken line).

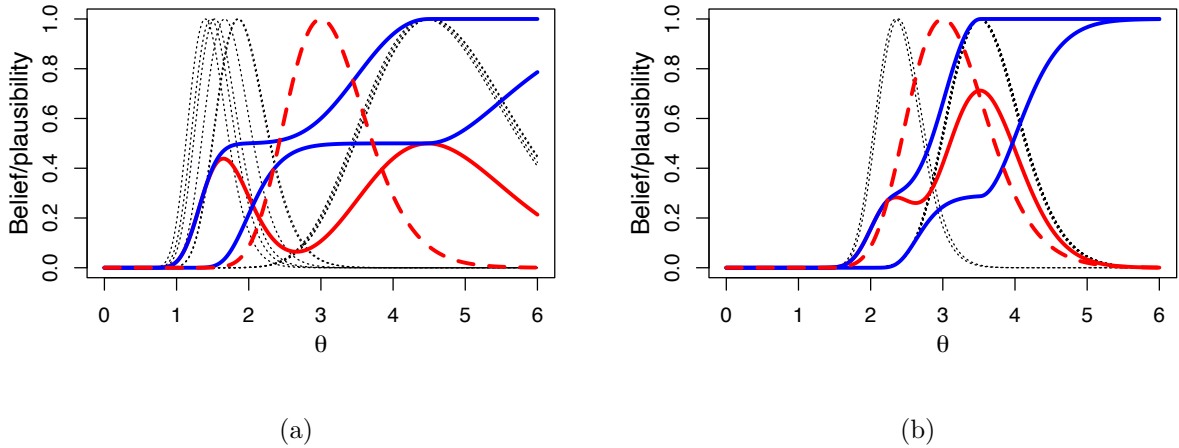


Figure 15: (a) Lognormal mt-GRFN prior $\tilde{\theta}_0 \sim 0.5T\tilde{N}(0.5, 0.1, 30, \log) + 0.5T\tilde{N}(1.5, 0.01, 20, \log)$ (contour function: red solid line, upper and lower cdfs: blues solid lines, and 10 realizations: black dotted lines), and lognormal approximation to the relative likelihood (red broken line); (b): mt-GRFN posterior $\tilde{\theta}'_x = \tilde{\theta}'_0 \oplus \tilde{L}_x$ (contour function: red solid line, upper and lower cdfs: blues solid lines, and 10 realizations: black dotted lines), and lognormal approximation to the relative likelihood (red broken line).

6. Conclusions

Until recently, the application of evidential reasoning to problems involving continuous variables has been limited due to the lack of practical models of belief functions in \mathbb{R}^p , $p \geq 1$ allowing for easy computation and combination. In [14], we have proposed a solution to this problem by considering “epistemic random fuzzy sets” as the basic construct, from which belief functions can be derived. In this framework, random fuzzy sets represent items of evidence and can be combined by a generalized product-intersection rule extending both the normalized product of possibility theory, and Dempster’s rule of Dempster-Shafer theory. In this framework, we have introduced in [14] simple models of random fuzzy sets based on the normal distribution. The proposed GRFN model and its multidimensional generalization define parameterized families of belief functions in \mathbb{R}^p that can be easily used in calculations and combined using simple mathematical formulas. However, GRFNs, indexed by three parameters (mean, standard deviation and precision) are not flexible enough to represent the wide variety of belief functions needed in applications.

In this paper, we have presented two extensions of the GRFN model. The first one consists in transforming a RFS \tilde{X} defined in \mathbb{R}^p by a bijective mapping ψ from \mathbb{R}^p to $\Lambda \subseteq \mathbb{R}^p$. Such a transformation allows us to define, e.g., belief functions on real intervals $[a, b]$ with $-\infty \leq a < b \leq +\infty$, or on the probability simplex \mathcal{S}_p . The second extension consists in considering mixtures of (transformed) GRFNs, which make it possible to define belief functions of arbitrary “shape” and complexity. The rich families of generalized GRFNs introduced in this paper are closed under the product-intersection rule and can be used for a variety of evidential reasoning tasks involving real variables.

We have discussed two important applications: the elicitation of expert beliefs about numerical quantities, and generalized Bayesian inference with weak priors defined as generalized GRFNs. Elicitation has been widely studied in the probabilistic context, much less in other settings such as possibility theory or belief functions. The new models introduced in this paper could be tested as representations of expert beliefs in real experiments. As for statistical inference, the epistemic random fuzzy set perspective provides a simple and consistent model in which data and prior knowledge are treated symmetrically as pieces of evidence represented by random fuzzy sets. The parametric families of random fuzzy numbers introduced in this paper can be used with a variety of statistical models without resorting to Monte Carlo simulation. It would be interesting to compare this approach with alternative methods such as, e.g., robust Bayesian analysis, both conceptually and practically.

As another important application of the models introduced in this paper, we can mention the quantification of prediction uncertainty in machine learning, for regression problems characterized by asymmetric or heavy-tailed noise distributions, or in which target variables are subject to some constraints. The EVREG model, an evidential neural network model proposed in [9, 11] for “classical” regression problems, models prediction uncertainty using GRFNs. It could be adapted, for instance, to compositional regression tasks (in which target variables are proportions) by transforming its outputs to logistic normal random fuzzy vectors. This and other research directions will be investigated in future publications.

693 **References**

- 694 [1] J. Aitchison and S. M. Shen. Logistic-normal distributions: Some properties and uses. *Biometrika*,
695 67(2):261–272, 1980.
- 696 [2] M. A. Aljarrah, C. Lee, and F. Famoye. On generating T-X family of distributions using quantile
697 functions. *Journal of Statistical Distributions and Applications*, 1(1):2, 2014.
- 698 [3] I. Couso and L. Sánchez. Higher order models for fuzzy random variables. *Fuzzy Sets and Systems*,
699 159(3):237–258, 2008.
- 700 [4] I. Couso and L. Sánchez. Upper and lower probabilities induced by a fuzzy random variable. *Fuzzy*
701 *Sets and Systems*, 165(1):1–23, 2011.
- 702 [5] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathe-*
703 *matical Statistics*, 38:325–339, 1967.
- 704 [6] T. Denœux. Inner and outer approximation of belief structures using a hierarchical clustering approach.
705 *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(4):437–460, 2001.
- 706 [7] T. Denœux. Likelihood-based belief function: justification and some extensions to low-quality data.
707 *International Journal of Approximate Reasoning*, 55(7):1535–1547, 2014.
- 708 [8] T. Denœux. Belief functions induced by random fuzzy sets: A general framework for representing
709 uncertain and fuzzy evidence. *Fuzzy Sets and Systems*, 424:63–91, 2021.
- 710 [9] T. Denœux. An evidential neural network model for regression based on random fuzzy numbers. In
711 S. Le Hégarat-Masclé, I. Bloch, and E. Aldea, editors, *Belief Functions: Theory and Applications*, pages
712 57–66, Cham, 2022. Springer International Publishing.
- 713 [10] T. Denœux. Belief functions on the real line defined by transformed Gaussian random fuzzy numbers.
714 In *2023 IEEE International Conference on Fuzzy Systems*, Songdo Incheon, Korea, August 2023. IEEE.
- 715 [11] T. Denœux. Quantifying Prediction Uncertainty in Regression using Random Fuzzy Sets: the ENNreg
716 model. *IEEE Transactions on Fuzzy Systems*, 2023. doi:10.1109/TFUZZ.2023.3268200.
- 717 [12] T. Denœux, D. Dubois, and H. Prade. Representations of uncertainty in artificial intelligence: Beyond
718 probability and possibility. In P. Marquis, O. Papini, and H. Prade, editors, *A Guided Tour of Artificial*
719 *Intelligence Research*, volume 1, chapter 4, pages 119–150. Springer Verlag, 2020.
- 720 [13] T. Denœux, D. Dubois, and H. Prade. Representations of uncertainty in artificial intelligence: Prob-
721 ability and possibility. In P. Marquis, O. Papini, and H. Prade, editors, *A Guided Tour of Artificial*
722 *Intelligence Research*, volume 1, chapter 3, pages 69–117. Springer Verlag, 2020.
- 723 [14] T. Denœux. Reasoning with fuzzy and uncertain evidence using epistemic random fuzzy sets: General
724 framework and practical models. *Fuzzy Sets and Systems*, 453:1–36, 2023.
- 725 [15] D. Dubois and H. Prade. *Possibility Theory: An approach to computerized processing of uncertainty*.
726 Plenum Press, New-York, 1988.
- 727 [16] D. Dubois, H. Prade, and R. Yager. Merging fuzzy information. In J. C. Bezdek, D. Dubois, and
728 H. Prade, editors, *Fuzzy sets in approximate reasoning and information systems*, pages 335–401. Kluwer
729 Academic Publishers, Boston, 1999.
- 730 [17] P. H. Garthwaite, J. B. Kadane, and A. O’Hagan. Statistical methods for eliciting probability distri-
731 butions. *Journal of the American Statistical Association*, 100(470):680–700, 2005.
- 732 [18] M. A. Gil, M. López-Díaz, and D. A. Ralescu. Overview on the development of fuzzy random variables.
733 *Fuzzy Sets and Systems*, 157(19):2546–2557, 2006.
- 734 [19] R. Haenni and N. Lehmann. Resource bounded and anytime approximation of belief function compu-
735 tations. *International Journal of Approximate Reasoning*, 31(1):103–154, 2002.
- 736 [20] N. L. Johnson. Systems of frequency curves generated by methods of translation. *Biometrika*,
737 36(1/2):149–176, 1949.
- 738 [21] R. Kruse and K. D. Meyer. *Statistics with vague data*. Kluwer, Dordrecht, 1987.
- 739 [22] J. D. Lowrance, T. D. Garvey, and T. M. Strat. A framework for evidential-reasoning systems. In T. K.
740 et al., editor, *Proceedings of AAAI’86*, volume 2, pages 896–903, Philadelphia, August 1986. AAAI.
- 741 [23] G. J. McLachlan and D. Peel. *Finite Mixture Models*. Wiley, New York, 2000.
- 742 [24] H. T. Nguyen. On random sets and belief functions. *Journal of Mathematical Analysis and Applications*,
743 65:531–542, 1978.

- 744 [25] M. L. Puri and D. A. Ralescu. Fuzzy random variables. *Journal of Mathematical Analysis and Appli-*
745 *cations*, 114(2):409–422, 1986.
- 746 [26] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.
- 747 [27] G. Shafer. Allocations of probability. *Annals of Probability*, 7(5):827–839, 1979.
- 748 [28] D. A. Sprott. *Statistical Inference in Science*. Springer-Verlag, Berlin, 2000.
- 749 [29] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning –I.
750 *Information Sciences*, 8:199–249, 1975.
- 751 [30] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.
- 752 [31] L. A. Zadeh. Fuzzy sets and information granularity. In M. M. Gupta, R. K. Ragade, and R. R. Yager,
753 editors, *Advances in Fuzzy Sets Theory and Applications*, pages 3–18. North-Holland, Amsterdam, 1979.

754 Appendix A. Combination of Gaussian Random Fuzzy vectors

755 Let $\tilde{X}_1 \sim \tilde{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \mathbf{H}_1)$ and $\tilde{X}_2 \sim \tilde{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \mathbf{H}_2)$ be two independent GRFVs such
756 that matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, \mathbf{H}_1 and \mathbf{H}_2 are all positive definite. We have

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

757 with

$$\mathbf{H}_{12} = \mathbf{H}_1 + \mathbf{H}_2, \quad \tilde{\boldsymbol{\mu}}_{12} = \mathbf{A}\tilde{\boldsymbol{\mu}}, \quad \text{and} \quad \tilde{\boldsymbol{\Sigma}}_{12} = \mathbf{A}\tilde{\boldsymbol{\Sigma}}\mathbf{A}^T,$$

758 where \mathbf{A} is the constant $p \times 2p$ matrix defined as

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

759

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

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

760 and

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

761 Furthermore, the degree of conflict between \tilde{X}_1 and \tilde{X}_2 is

$$\kappa = 1 - \sqrt{\frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}} \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\mu}_2 - \tilde{\boldsymbol{\mu}}^T \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{\boldsymbol{\mu}} \right] \right\}. \quad (\text{A.1})$$

762 Appendix B. Proof of Proposition 3

763 The membership function of $\tilde{Y}(\omega) \sim T\tilde{N}(\mu, \sigma^2, h, \log)$ is

$$\tilde{Y}(\omega)(y) = \exp \left(-\frac{h}{2} (\log y - M(\omega))^2 \right),$$

764 with $M \sim N(\mu, \sigma^2)$. The α -cut of $\tilde{Y}(\omega)$ is

$${}^\alpha\tilde{Y}(\omega) = \left[\exp\left(M(\omega) - \sqrt{\frac{-2\log\alpha}{h}}\right), \exp\left(M(\omega) + \sqrt{\frac{-2\log\alpha}{h}}\right) \right]. \quad (\text{B.1})$$

765 Now, $M - \sqrt{\frac{-2\log\alpha}{h}} \sim N\left(\mu - \sqrt{\frac{-2\log\alpha}{h}}, \sigma^2\right)$, hence the lower bound of (B.1) has a lognormal
766 distribution $LN\left(\mu - \sqrt{\frac{-2\log\alpha}{h}}, \sigma^2\right)$ and an expectation

$$\mathbb{E}_*({}^\alpha\tilde{Y}) = \exp\left(\mu - \sqrt{\frac{-2\log\alpha}{h}} + \frac{\sigma^2}{2}\right). \quad (\text{B.2})$$

767 Similarly,

$$\mathbb{E}^*({}^\alpha\tilde{Y}) = \exp\left(\mu + \sqrt{\frac{-2\log\alpha}{h}} + \frac{\sigma^2}{2}\right). \quad (\text{B.3})$$

768 The lower and upper expectations of \tilde{Y} are, respectively, the integrals of (B.2) and (B.3)
769 from $\alpha = 0$ to $\alpha = 1$. Let us start with

$$\mathbb{E}_*(\tilde{Y}) = \int_0^1 \exp\left(\mu - \sqrt{\frac{-2\log\alpha}{h}} + \frac{\sigma^2}{2}\right) d\alpha.$$

770 By the change of variable $\beta = \sqrt{\frac{-2\log\alpha}{h}}$, we get

$$\mathbb{E}_*(\tilde{Y}) = h \exp\left(\mu + \frac{\sigma^2}{2}\right) \int_0^{+\infty} \beta \exp\left(-\frac{h}{2}\beta^2 - \beta\right) d\beta.$$

771 Completing the square gives us

$$\mathbb{E}_*(\tilde{Y}) = h \exp\left(\mu + \frac{\sigma^2}{2} + \frac{1}{2h}\right) \underbrace{\int_0^{+\infty} \beta \exp\left[-\frac{h}{2}\left(\beta + \frac{1}{h}\right)^2\right] d\beta}_I. \quad (\text{B.4})$$

772 Integral I is related to the mean of the left-truncated normal distribution on $[0, +\infty)$, with
773 mean $-1/h$, standard deviation $1/h$, and density

$$f(y) = \frac{\sqrt{h}}{\sqrt{2\pi}\left(1 - \Phi\left(\frac{1}{\sqrt{h}}\right)\right)} \exp\left[-\frac{h}{2}\left(\beta + \frac{1}{h}\right)^2\right].$$

774 The mean of this truncated normal distribution is

$$-\frac{1}{h} + \frac{1}{\sqrt{h}} \frac{\phi(1/\sqrt{h})}{1 - \Phi(1/\sqrt{h})},$$

775 where ϕ is the standard normal pdf. Consequently, we have

$$I = \frac{-\frac{1}{h} + \frac{1}{\sqrt{h}} \frac{\phi(1/\sqrt{h})}{1 - \Phi(1/\sqrt{h})}}{\frac{\sqrt{h}}{\sqrt{2\pi}(1 - \Phi(\frac{1}{\sqrt{h}}))}} = \sqrt{2\pi} \left[-h^{-3/2} \left(1 - \Phi \left(\frac{1}{\sqrt{h}} \right) \right) + \frac{1}{h} \phi \left(\frac{1}{\sqrt{h}} \right) \right]. \quad (\text{B.5})$$

776 From (B.4) and (B.5), we get (17a). Similar calculations starting from the upper bound of
777 (B.1) yield (17b).

778 Appendix C. Proof of Theorem 4

779 We start by the following lemma, which generalizes Lemma 1 in [14].

780 **Lemma 3.** *Let (M_1, Z_1) and (M_2, Z_2) be two independent two-dimensional random vectors
781 such that*

$$P(Z_1 = k) = \pi_{1k}, \quad k = 1, \dots, K,$$

$$M_1 | (Z_1 = k) \sim N(\mu_{1k}, \sigma_{1k}^2), \quad k = 1, \dots, K$$

783 and

$$P(Z_2 = \ell) = \pi_{2\ell}, \quad \ell = 1, \dots, L,$$

$$M_2 | (Z_2 = \ell) \sim N(\mu_{2\ell}, \sigma_{2\ell}^2), \quad \ell = 1, \dots, L.$$

Let \tilde{F} be the fuzzy subset of $\mathbb{R}^2 \times \{1, \dots, K\} \times \{1, \dots, L\}$ with membership function

$$\begin{aligned} \tilde{F}(m_1, m_2, y_1, y_2) &= \text{hgt}(\text{GFN}(m_1, h_1(y_1)) \cdot \text{GFN}(m_2, h_2(y_2))) \\ &= \exp \left(-\frac{h_1(y_1)h_2(y_2)(m_1 - m_2)^2}{2(h_1(y_1) + h_2(y_2))} \right), \end{aligned}$$

785 where $h_1(y_1) = \prod_{k=1}^K h_{1k}^{y_{1k}}$ and $h_2(y_2) = \prod_{\ell=1}^L h_{2\ell}^{y_{2\ell}}$ and, as before, $y_{1k} = I(y_1 = k)$ and
786 $y_{2\ell} = I(y_2 = \ell)$.

787 The conditional probability distribution of (M_1, M_2, Z_1, Z_2) given \tilde{F} can be described as
788 follows:

- 789 • The conditional probability distribution of (M_1, M_2) given \tilde{F} and $(Z_1, Z_2) = (k, \ell)$ is
790 two-dimensional Gaussian with mean vector $\tilde{\boldsymbol{\mu}}_{k\ell} = (\tilde{\mu}_{1k\ell}, \tilde{\mu}_{2k\ell})^T$ and covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{k\ell} = \begin{pmatrix} \tilde{\sigma}_{1k\ell}^2 & \rho_{k\ell} \tilde{\sigma}_{1k\ell} \tilde{\sigma}_{2k\ell} \\ \rho_{k\ell} \tilde{\sigma}_{1k\ell} \tilde{\sigma}_{2k\ell} & \tilde{\sigma}_{2k\ell}^2 \end{pmatrix},$$

with

$$\tilde{\mu}_{1k\ell} = \frac{\mu_{1k}(1 + \bar{h}\sigma_{2\ell}^2) + \mu_{2\ell}\bar{h}_{k\ell}\sigma_{1k}^2}{1 + \bar{h}_{k\ell}(\sigma_{1k}^2 + \sigma_{2\ell}^2)}, \quad \tilde{\mu}_{2k\ell} = \frac{\mu_{2\ell}(1 + \bar{h}_{k\ell}\sigma_{1k}^2) + \mu_{1k}\bar{h}_{k\ell}\sigma_{2\ell}^2}{1 + \bar{h}_{k\ell}(\sigma_{1k}^2 + \sigma_{2\ell}^2)} \quad (\text{C.1a})$$

$$\tilde{\sigma}_{1k\ell}^2 = \frac{\sigma_{1k}^2(1 + \bar{h}_{k\ell}\sigma_{2\ell}^2)}{1 + \bar{h}_{k\ell}(\sigma_{1k}^2 + \sigma_{2\ell}^2)}, \quad \tilde{\sigma}_{2k\ell}^2 = \frac{\sigma_{2\ell}^2(1 + \bar{h}_{k\ell}\sigma_{1k}^2)}{1 + \bar{h}_{k\ell}(\sigma_{1k}^2 + \sigma_{2\ell}^2)} \quad (\text{C.1b})$$

$$\rho_{k\ell} = \frac{\bar{h}_{k\ell}\sigma_{1k}\sigma_{2\ell}}{\sqrt{(1 + \bar{h}_{k\ell}\sigma_{1k}^2)(1 + \bar{h}_{k\ell}\sigma_{2\ell}^2)}}, \quad (\text{C.1c})$$

791 where

$$\bar{h}_{k\ell} = \frac{h_{1k}h_{2\ell}}{h_{1k} + h_{2\ell}}. \quad (\text{C.1d})$$

792 • The conditional probability distribution of (Z_1, Z_2) given \tilde{F} is

$$P(Z_1 = k, Z_2 = \ell \mid \tilde{F}) = \tilde{\pi}_{k\ell} = \frac{(1 - \kappa_{k\ell})\pi_{1k}\pi_{2\ell}}{\sum_{k', \ell'} (1 - \kappa_{k'\ell'})\pi_{1k'}\pi_{2\ell'}}, \quad (\text{C.2})$$

793 where $\kappa_{k\ell}$ is the degree of conflict between two independent GRFNs $\tilde{X}_{1k} \sim \tilde{N}(\mu_{1k}, \sigma_{1k}^2, h_{1k})$
794 and $\tilde{X}_{2\ell} \sim \tilde{N}(\mu_{2\ell}, \sigma_{2\ell}^2, h_{2\ell})$ given by Proposition 1.

Proof. Given (Z_1, Z_2) , M_1 and M_2 have normal distributions. The conditional probability distribution of (M_1, M_2) given \tilde{F} and $(Z_1, Z_2) = (k, \ell)$ results directly from Lemma 1 of [14]. Now, from Bayes' theorem,

$$\begin{aligned} P(Z_1 = k, Z_2 = \ell \mid \tilde{F}) &= \frac{P(\tilde{F} \mid Z_1 = k, Z_2 = \ell)P(Z_1 = k, Z_2 = \ell)}{P(\tilde{F})} \\ &= \frac{P(\tilde{F} \mid Z_1 = k, Z_2 = \ell)\pi_{1k}\pi_{2\ell}}{\sum_{k', \ell'} P(\tilde{F} \mid Z_1 = k', Z_2 = \ell')\pi_{1k'}\pi_{2\ell'}}. \end{aligned}$$

795 From Lemma 1 in [14], $P(\tilde{F} \mid Z_1 = k, Z_2 = \ell) = 1 - \kappa_{k\ell}$, where $\kappa_{k\ell}$ is given by Proposition
796 1, which completes the proof. \square

797 We can now give the proof of Theorem 4.

798 *Proof of Theorem 4.* Let (M_1, Z_1) and (M_2, Z_2) be pairs of random variables from $(\Omega_1, \Sigma_1, P_1)$
799 and $(\Omega_2, \Sigma_2, P_2)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ corresponding, respectively, to m-GRFNs \tilde{X}_1 and \tilde{X}_2 . The
800 orthogonal sum of \tilde{X}_1 and \tilde{X}_2 is the RFS $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \tilde{P}_{12}, \mathbb{R}, \mathcal{B}(\mathbb{R}), \tilde{X}_\odot)$, where \tilde{X}_\odot is
801 the mapping

$$\tilde{X}_\odot : (\omega_1, \omega_2) \rightarrow \text{GFN}(M_{12}(\omega_1, \omega_2), h_1(\omega) + h_2(\omega)),$$

802 with

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

and \tilde{P}_{12} is the probability measure on $\Omega_1 \times \Omega_2$ obtained by conditioning $P_1 \times P_2$ on the fuzzy set $\tilde{\Theta}^*(\omega_1, \omega_2) = \text{hgt}(\text{GFN}(M_1(\omega_1), h_1(\omega)), \text{GFN}(M_2(\omega_2), h_2(\omega)))$. The pushforward measure of \tilde{P}_{12} by the random vector (M_1, M_2, Z_1, Z_2) is the conditional distribution given \tilde{F} described in Lemma 3, with parameters $(\tilde{\mu}_{1k\ell}, \tilde{\mu}_{2k\ell}, \tilde{\sigma}_{1k\ell}, \tilde{\sigma}_{2k\ell}, \rho_{k\ell}, \tilde{\pi}_{j\ell})$, $k = 1, \dots, K$, $\ell = 1, \dots, L$. The conditional expectation of M_{12} given $(Z_1, Z_2) = (k, \ell)$ and \tilde{F} is

$$\begin{aligned} \mathbb{E}(M_{12} \mid Z_1 = k, Z_2 = \ell, \tilde{F}) &= \frac{h_{1k}\mathbb{E}(M_1 \mid Z_1 = k, Z_2 = \ell, \tilde{F}) + h_{2\ell}\mathbb{E}(M_2 \mid Z_1 = k, Z_2 = \ell, \tilde{F})}{h_{1k} + h_{2\ell}} \\ &= \frac{h_{1k}\tilde{\mu}_{1k} + h_{2\ell}\tilde{\mu}_{2\ell}}{h_{1k} + h_{2\ell}}, \end{aligned}$$

and its conditional variance is

$$\begin{aligned} \text{Var}(M_{12}|Z_1 = k, Z_2 = \ell, \tilde{F}) &= \frac{1}{(h_{1k} + h_{2\ell})^2} (h_1^2 \text{Var}(M_1|Z_1 = k, Z_2 = \ell, \tilde{F}) + \\ &h_2^2 \text{Var}(M_2|Z_1 = k, Z_2 = \ell, \tilde{F}) + 2h_{1k}h_{2\ell} \text{Cov}(M_1, M_2|Z_1 = k, Z_2 = \ell, \tilde{F})), \end{aligned}$$

803 which gives

$$\text{Var}(M_{12}|Z_1 = k, Z_2 = \ell, \tilde{F}) = \frac{h_{1k}^2 \tilde{\sigma}_{1k}^2 + h_{2\ell}^2 \tilde{\sigma}_{2\ell}^2 + 2\rho_{k\ell} h_{1k} h_{2\ell} \tilde{\sigma}_{1k} \tilde{\sigma}_{2\ell}}{(h_{1k} + h_{2\ell})^2}.$$

Finally, $P(Z_1 = k, Z_2 = \ell | \tilde{F}) = \tilde{\pi}_{k\ell}$, which completes the proof of the first part of the theorem. The second part is obtained directly by noticing that

$$\begin{aligned} \kappa &= 1 - P(\tilde{F}) = 1 - \sum_{k=1}^K \sum_{\ell=1}^L P(\tilde{F} | Z_1 = k, Z_2 = \ell) P(Z_1 = k, Z_2 = \ell) \\ &= 1 - \sum_{k=1}^K \sum_{\ell=1}^L (1 - \kappa_{k\ell}) \pi_{1k} \pi_{2\ell} = \sum_{k=1}^K \sum_{\ell=1}^L \kappa_{k\ell} \pi_{1k} \pi_{2\ell}. \end{aligned}$$