On the negation intensity of a probability distribution

Xiangjun Mi[a], Chongru Huang[a], Bingyi Kang[a,b,*]

[a] *College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, 712100, China.*
[b] *Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China.*
[c] *Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China.*

Abstract

How to obtain negation knowledge is a crucial topic, especially in the field of artificial intelligence. Limited work has been done on the negation of a probability distribution and has been studied in depth throughout the literature. However, the aspect of the intensity level of negation enforcement has not yet been investigated. Moreover, let us note that the main characteristic of intelligent systems is just the flexibility for the sake of being able to represent knowledge according to each situation. In general, researchers have a tendency to express the need for cognitive range in the negation. Thus, it would seem very useful to find a wide range of negations under intensity levels in a probability distribution. Based on these ideas, this paper first proposes a new approach of finding a probability distribution negation and gives a domain of intensity in which the negation is executed, which is called the negation space. Then, we investigate a number of desirable properties and explore their correlation with entropy. Numerical examples show the characteristics of the proposed negation solution. Finally, we validate the efficiency of the proposed method from the point of view of the Dempster-Shafer belief structure.

Keywords: Negation, probability distributions, aggregation, decision-making, membership grade, uncertainty.

*Corresponding author: B. Kang, College of Information Engineering, Northwest A&F University, Yangling, Shaanxi, 712100, China. Email address: bingyi.kang@nwsuaf.edu.cn; bingyi.kang@hotmail.com.

1. Introduction

In any case, the representation of knowledge is an extremely important issue for information science, especially in the construction of artificial intelligence. A large number of studies have been developed to solve the issue of information characterisation of knowledge contained in information sources, such as fuzzy sets [1, 2, 3], Dempster-Shafer evidence theory [4, 5, 6, 7], Z numbers [8, 9, 10], evidence reasoning [11], etc. They have been accepted as the driving force behind the move from theory to practice, and are widely used in decision making [12, 13], information fusion [14, 15, 16], and intelligent systems [17, 18].

As rare events can in some cases have a serious impact on the system, the study of rare events is essential [19]. Negation provides us a new way of thinking about problem solving. For instance, "What is it?" This may not be directly answerable in some contexts. One angle we can choose to address is "What is its negation?". Furthermore, when a theorem is difficult to prove directly, we can easily prove it wrong if we can find a counter-example of it. Just as in the original Dempster-Shafer theory [20, 21], researchers were unable to prove that the Dempster's combination rule is irrational by strict mathematical deduction. But since Zadeh found a counter-example [22], it has been confirmed that using Dempster's combination rule to combine highly conflicting evidence yields results that defy human intuition. Thus, this gives us sufficient grounds to justify the shortcomings of the original combination rule. Here, let us consider a question in a knowledge-based rule base. If A is *good*, then B is p. If A is *bad*, then A is q. If we represent *good* as a fuzzy set, the process of becoming *bad* is well known. But if we use a probability distribution to determine the concept of *good*, the determination *bad* becomes the negation of finding the probability distribution. In this paper, we are concerned with the form of knowledge contained in the negation of a probability distribution.

Formulating the negation of the probability event was first formally put forward by Zadeh in his BISC blog. Since then, the topic has become more and more popular. Yager [23] presented a probability distribution of the negation procedure from the point of view of maximum entropy. In his negation model, each probability is treated as a separate element x_i and its complementary probabilities are characterised using $1 - p_i$, and finally all complementary probabilities are normalised with $n-1$ to obtain the negation of the probability distribution. Yager's idea of design negation can be summarised as follows: if this probability distribution is initially certain information, then the iterations of its negation will eventually reach a state of maximum entropy, i.e., a state of complete uncertainty. Subsequently, the negation model based on maximum entropy has attracted widespread attention from scholars, who have extended the model to areas such as evidence theory [24, 25, 26, 27, 28], Z-numbers [29] and have achieved good performance in the characterisation and treatment of uncertain information in these areas. Moreover, following the negation model based on maximum entropy proposed by Yager, some *Bayesian* properties [30], extended algorithms [31, 32], entropy measures [33] and complex-valued distribution [34] have recently been studied by many scholars. These studies show that the subject of obtaining a probability distribution negation method is an open and important topic.

The main goals of this paper is to find a more general method for obtaining the negation of a probability distributions, and to do so in a way that is consistent with our cognition and intuition. More specifically, the main contributions of this paper are summarised below. First, we suggest a negation procedure to obtain the negation of a probability distribution based on the idea of order, and formulate a domain of negation intensity, which we call the negation space, by which we quantify the intensity level of negation execution. Then we investigate the properties of the negation method from the perspective of entropy. Finally, we show that the proposed negation scheme is indeed a desirable negation in terms of the conflict measurement of belief structure in Dempster-Shafer evidence theory.

2. Negation of a probability distribution

Here we would like to focus on the concern about how to find a negation of a probability distribution. In the following, we shall suggest a negation of a probability distribution.

Assume that one of the frames we refer to is a set $X = \{x_1, ..., x_i, ..., x_n\}$. Consider a probability distribution over X denoted as

$$P = \{p_1, ..., p_i, ..., p_n\}.$$

Naturally, we have

$$\sum_{i=1}^{n} p_i = 1 \ and \ 0 \leq p_i \leq 1, \forall i = 1, ..., n.$$

Note that in this paper, we stipulate that the negation of the probability distribution P is expressed as $\bar{P} = \{\bar{p}_1, ..., \bar{p}_i, ..., \bar{p}_n\}$. In effect, the negation grants us a statement about representational knowledge that allows us to use "*not P*".

Below we shall introduce in detail our motivation in constructing a negation scheme. Before proceeding to our task, defining the negation of the probability distribution, we shall find it convenient to use a slightly different notation for the proposed negation method. Hence, we first make the following setting. For the probability distribution P, if the size of p_i is ordered ascending, correspondingly, we can obtain a set of sequences about x_i. In order to find this sequence, we let δ_i be an index function, which makes $\delta_i(k)$ the kth most probability index associated with candidate x_i in X. Thus here $p_{\delta_i(k)}$ is the kth largest probability associated with x_i in P. And we let them satisfy $p_{\delta_i(1)} \leq \cdots \leq p_{\delta_i(k)} \leq \cdots \leq p_{\delta_i(n)}$. Based on the idea of negation, we can easily derive an intuitive result that $\bar{p}_{\delta_i(1)} \geq \cdots \geq \bar{p}_{\delta_i(k)} \geq \cdots \geq \bar{p}_{\delta_i(n)}$.

Now, we define an aggregate-type operator, which we call *joint income* (JI). This operator is used to collect a set of values and provide a single value. We define this operator as follows

$$JI_z(p_{\delta_i(1)}, ..., p_{\delta_i(n)}) = \sum_{j=1}^{z}\left(1 + \sum_{\substack{j=1, j\neq z;\\ k=1}}^{n} Inc(p_{\delta_z(k)}, p_{\delta_j(k)})\right).$$

We shall denote $Inc(p, q)$ as the income for p from q, which is the degree of support received. Generally, we assume that $Inc(p, q)$ satisfies the following three desired properties:

- $Inc(p, q) \in [0, 1]$;
- $Inc(p, q) = Inc(q, p)$;
- $Inc(p, q) \geq Inc(x, y),\ if\ |p - q| < |x - y|$.

Thus we see the more similar, the closer two values, the more they support each other. Here, a possible form of the Inc function is provided as

$$Inc(p, q) = (1 - |p - q|).$$

Obviously, it is easy to deduce that the Inc function given above satisfies the three important properties mentioned earlier.

Next, base on the JI operator, we shall define a unique operator, we call the *power negation* (PN). But before that, we need to determine a general regular increasing monotone function. Let $U : [0, 1] \to [0, 1]$, it satisfies the constraint: $U(0) = 0$; $U(1) = 1$; and $U(x) \geq U(y)$ if $x > y$. By appropriately selecting U, we can implement different types of aggregation imperative. Here, we draw up a simple function that satisfies the above conditions as $f(x) = x^{\upsilon}$ with $\upsilon \geq 0$. Thus this operator is formally defined by

$$PN_z(p_{\delta_i(1)}, ..., p_{\delta_i(n)}) = \left(\frac{JI_z}{JI_n}\right)^{1-\kappa} - \left(\frac{JI_{z-1}}{JI_n}\right)^{1-\kappa}$$

where κ is defined as the negation parameter. We call $\kappa \in [0, 1]$ the negation space, which is used to characterize the possible range of a probability distribution after it has been negated in a single negation operation. In fact, the magnitude of the value of κ indicates the strength of negation being executed. The higher κ is, the stronger it is. Typically, κ takes the value 0, 0.1, ..., 1. We call $\kappa = 0$ the lower bound of negation and $\kappa = 1$ the upper bound of negation. Note that for the convenience of presentation, this paper records the negation of the probability distribution P at the lower bound of negation as $\bar{P}_*$, and its component element is represented as $\bar{p}_{*i}$. Correspondingly, the negation it obtains at the upper bound of negation as $\bar{P}^*$, and its component element is represented as $\bar{p}_i^*$. We shall explain later in detail the implications of the suggested negation space for the negation of a probability distribution. In addition, we note that when the minimum value of z is equal to 1, $z-1$ crosses the boundary, making JI_{z-1} invalid. Thus, by defining $JI_{z-1} = 0$, we end up with a complete PN_z.

In order to fusion the probability information generated after using the PN operator, we expect to use this set of output values to correspondingly multiply each probability in P. However, we note that there may be values of zero for these probabilities in P, rendering some output of the PN operator unusable. For this reason, we introduce a monotonically decreasing function, denoted as $g(x) = e^{-x}$. Thus, we have

$$Prod_i = PN_i \times e^{-p_{\delta_i(k)}}.$$

It is interesting to note, however, that the output of $Prod_i$ may be a set of subsequences whose sum is not equal to 1, so that a necessary remedial step before finally obtaining this sequence is to perform a normalization operation on it. Thus, we can obtain the sequence of the probability distribution after negation as

$$\bar{p}_{\delta_i(k)} = \frac{Prod_i}{\sum\limits_{i=1}^{n} Prod_i}.$$

According to the index function $\delta_i(k)$, correspondingly, we assign $\bar{p}_{\delta_i(k)}$ to the candidate x_i. Thus, in this case

$$\bar{P}: \ \bar{p}_{\delta_i(k)} \overset{refer\ to\ \delta_i(k)}{\rightarrow} \bar{p}_i.$$

In this way, we end up with the negation of the probability distribution.

In fact, the proposed negation method first finds the ascending sequence of all probabilities in a probability distribution through a special index function, and uses this index function to record their corresponding positions with the elements in the reference frame X. The realization process of the negation operation of the probability distribution is actually to gradually reduce the value of the elements with high probability in this group of sequences, and gradually increase the value of the elements with low probability. Finally, the negated probability is assigned to the corresponding candidate element in X by using the index function. The final result of the negation of the probability distribution depends on the size of the negation parameter in the negation space.

In what follows, we shall further examine some interesting properties about the negation of a probability distribution, but before we do so, we first need to demonstrate that $\bar{P}$ is also a probability distribution.

THEOREM 1. *Assume that for a probability distribution $P = \{p_1, ..., p_i, ..., p_n\}$, its negation is given through our proposed scheme, denoted by $\bar{P}$. Then, for all i, we have*

$$\sum_{i=1}^{n} \bar{p}_i = 1 \ and \ 0 \leq \bar{p}_i \leq 1.$$

PROOF. *Recalling our previous definitions, it is easy to derive that $Inc(p_{\delta_z(k)}, p_{\delta_j(k)}) > 0$, and $JI_z(p_{\delta_i(1)}, ..., p_{\delta_i(n)}) > 0$. For PN_i, as previously stated*

$$PN_z = \left(\frac{JI_z}{JI_n}\right)^{\upsilon} - \left(\frac{JI_{z-1}}{JI_n}\right)^{\upsilon}.$$

Since $f(x) = x^{\upsilon}$ is a monotonic increasing function with $\upsilon \in [0, 1]$, we can say that $PN_z > 0$ as long as $JI_z > JI_{z-1}$. In the following, we will derive it step *by step. First, we know $1 \leq z \leq n$. When $z = 1$, we have*

$$PN_1 = \left(\frac{JI_1}{JI_n}\right)^{\upsilon} - \left(\frac{JI_0}{JI_n}\right)^{\upsilon}.$$

Due to $JI_1 > 0$ and $JI_0 = 0$, we have $PN_1 > 0$. When $1 < z \leq n$, we have $JI_z - JI_{z-1} > 0$. Thus, we can conclude $JI_z > 0$. Futhermore, if we let $\kappa = 1$, then $\upsilon = 0$. For $\bar{p}_{\delta_i(n)}$, we have $\bar{p}_i = 1$, since $PN_1 = 1$. In addition to this case, it is easy to infer $\bar{p}_i \geq 0$. Thus, we can draw the first conclusion $\bar{p}_i \in [0, 1]$.

Then, let us consider the next question. As in the preceding

$$\bar{p}_{\delta_i(k)} = \frac{Prod_i}{\sum\limits_{i=1}^{n} Prod_i}.$$

We know that the index function $\delta_i(k)$ only changes the probability and the position it corresponds to some random variable. Therefore, it is easy to infer that $\sum_{i=1}^{n} \bar{p}_{\delta_i(k)} \Leftrightarrow \sum_{i=1}^{n} \bar{p}_i$. Then, we have

$$\begin{aligned}\sum_{i=1}^{n} \bar{p}_i &= \sum_{i=1}^{n} \bar{p}_{\delta_i(k)} \\ &= \sum_{i=1}^{n} \left(\frac{Prod_i}{\sum_{i=1}^{n} Prod_i}\right) \\ &= 1.\end{aligned}$$

Now, let us formally return to that place of interest. Recall that we define the range of possible values of the negation operation, which we call the negation space. Until then, it had been mysteriously defined. In the following, through some theorems, we shall explore more interesting findings about the negation parameter in the proposed negation method. One direction of exploration that needs to be pointed out in advance is that we try to establish the relationship between the level of negation and uncertainty. Before that, it is worth noting that there are many different measures of entropy we shall measure the uncertainty of a probability distribution as

$$H(P) = \sum_{i=1}^{n}(1-p_i)p_i = 1 - \sum_{i=1}^{n} p_i^2.$$

Note that here we prefer this form of entropy measure instead of the classical Shannon measure [35] because of its simplicity, i.e., the computation it entails, without the *log*. Since we do not have the additivity of entropy that requires independent probabilities, a unique property of Shannon entropy, we do not lose anything by using this measure.

THEOREM 2. *Let a probability distribution be $P = \{p_1, ..., p_n\}$. $\bar{P}_v$ denotes the negation of the probability distribution when the negation parameter takes the value of v, and v usually takes 0, 0.1, ..., 1. Then, we have*

$$H(\bar{P}_r) > H(\bar{P}_e) \text{ with } 0 \leq e, r \leq 1$$

where $e = r + 0.1$.

PROOF. *As noted earlier, we know $JI_z \leq JI_n$, so*

$$0 < \frac{JI_z}{JI_n} \leq 1.$$

Since $0 \leq \kappa \leq 1$, we know $0 \leq 1-\kappa \leq 1$. Thus $(JI_z/JI_n)^{1-\kappa}$ is a monotonically increasing function. Additionally, we can see that $e^{-p_{\delta_i(k)}}$ is also a monotone decreasing function. Due to the index function $\delta_i(k)$, the probability distribution is a set of ascending sequences, so $PN_i \times e^{-p_{\delta_i(k)}}$ is an increasing function. Grasping the properties of the above functions, we shall now discuss the role of the negative parameter taking a value.

On the one hand, when $\kappa = 0$. For PN_1, because of $JI_0 = 0$, we have

$$PN_1 = \frac{JI_1}{JI_n}.$$

Furthermore, if $z \geq 2$, in this case

$$PN_z = \frac{JI_z - JI_{z-1}}{JI_n}.$$

Moreover, the joint income function Inc causes JI_z to increase by a nearly equal gap which is approximately 1/n. This makes $\bar{P}_*$ reach the maximum entropy in the negation space.

On the other hand, when $\kappa = 1$, in this case

$$PN_1 = \left(\frac{JI_z}{JI_n}\right)^{1-1} = 1.$$

Since $\bar{P}$ also is a probability distribution, it makes $H(\bar{P}^*)$ minimum in the negation space. Hence we have

$$\begin{aligned} H(\bar{P}^*)_{min} &= \sum_{i=1}^{n}(1 - \bar{p}^*_{\delta_i(1)})\bar{p}^*_{\delta_i(1)} \\ &= 1 - \sum_{i=1}^{n}(\bar{p}^*_{\delta_i(1)})^2 \\ &= 1 - \sum_{i=1}^{n} 1^2 = 0. \end{aligned}$$

In addition to these two cases, let us consider the more general case. As κ changes from 0 to 1, the uncertainty of the $\bar{P}$ gets lower and lower, while PN_1 gets larger and larger, until it converges to 1, i.e., the minimum entropy.

Notably, in Theorem 2, we claim that the maximum value of the entropy of $\bar{P}^*$ is obtained in the negation space compared to $\bar{P}$ under other negation strength levels, so it is not really the maximum entropy that can be achieved under any conditions. As a matter of fact, in the proposed negation scheme, there is only one maximum entropy for negation of a probability distribution. That is, the negation of a uniform distribution is uniform. Below, we give an explanation.

THEOREM 3. *Assume $P = \{p_1, ..., p_i, ..., p_n\}$ is such that $p_i = 1/n$ for all i then*

$\bar{p}_i = 1/n$ for all i if and only if it is the lower bound of negation in the negation space. In this case, $\bar{P}$ achieves the maximum entropy, i.e., $H(\bar{P}_*)_{max}$.

PROOF. *As before, when $\kappa = 0$ and $z = 1$, we have*

$$PN_1 = \frac{JI_1}{JI_n}.$$

If $z \geq 2$, then

$$PN_z = \frac{JI_z - JI_{z-1}}{JI_n}.$$

Since $Inc(1/n, 1/n) = (1 - |1/n - 1/n|) = 1$, then $JI_1 = 1/n$ and hence

$$PN_z = \frac{JI_z - JI_{z-1}}{JI_n} = \frac{1}{n} \text{ with } z \geq 2.$$

In the situation where $1 \leq z \leq n$, since $Prod_i = PN_i \times e^{-(1/n)}$, we have

$$Prod_1 = \cdots = Prod_i = \cdots = Prod_n.$$

This gives us

$$\bar{p}_{\delta_i(k)} = \frac{Prod_i}{\sum_{i=1}^{n} Prod_i} = \frac{1}{n}.$$

We notice $p_i = 1/n$, so there is no need to use the index function $\delta_i(k)$ to mark the position of p_i and x_i in P. Thus, we now can obtain

$$\bar{p}_i = \frac{1}{n}.$$

Finally, in this particular situation, we can get

$$
\begin{aligned}
H(\bar{P}_*)_{max} &= \sum_{i=1}^{n}(1-\bar{p}_{\delta_i(k)})\bar{p}_{\delta_i(k)} \\
&= 1-\sum_{i=1}^{n}\bar{p}_{\delta_i(k)}^2 \\
&= 1-\left((\frac{1}{n})^2+\cdots+(\frac{1}{n})^2\right) \\
&= \frac{n-1}{n}.
\end{aligned}
$$

Based on Theorem 3, we can easily derive the following theorem.

THEOREM 4. *Assume P is a probability distribution, and $\bar{P}$ is its negation, then $H(\bar{P}_*) \geq H(P)$.*

PROOF. *The proof is obvious and trivial.*

Through the above theorem, we find that the negation space is an important domain space for obtaining the negation of a probability distribution. As the strength of the negation increases, the uncertainty of the probability distribution after negation becomes smaller and smaller, until it is equal to zero, which is $H(\bar{P}^*)_{min}$. One reason for this phenomenon is that the entropy of $\bar{P}$ is gradually decreasing. In other words, we can always find the opposite of a probability distribution, because the final state is the least uncertain. As mentioned before, in the proposed negation technique, we treat P as a whole, that is, a set of ascending sequences. Its negation is to increase the probability of candidates supported by some smaller probabilities in the set of sequences, until a certain candidate is completely affirmed, namely $\bar{p}_{\delta_i(1)} = 1$. At the same time, it makes the probability of some candidates that are supported by a larger probability smaller, until a certain candidate is completely rejected, that is, there is $\bar{p}_{\delta_i(n)} = 0$. When the negation strength is the upper bound of negation, the negation of the probability distribution is the strongest at this time, that is, the most unlikely event becomes the most likely event. Usually for researchers, getting the negation domain of an event is more important than its only opposite, which provides space for expressing cognition. Therefore, this discovery is intuitive and what we desire.

Furthermore, we completely analyze our idea of designing negation method. When the negation parameter is the lower bound of negation, the reason for choosing the near-maximum entropy alternative is that it selects an allowable alternative, which brings less unsupported information. In other words, in this case, the information we can use is not helpful because it has the greatest uncertainty and cannot make decisions. In addition, let us analogize to the real world and explain this connotation more deeply. Assuming that resources can be obtained equally, a phenomenon is that a few people often control most of the resources. The reason for this phenomenon may be caused by the uneven distribution of initial resources. It is well known that people with sound intelligence are more likely to be successful than people with disabilities. The negation in this case is the final negation after considering the time span. Although we do not clearly point out whether this non-uniform probability distribution reached a certain fixed asymmetric distribution after negation, in this case, at least it is not the state of maximum entropy. In other words, the negation of the asymmetric probability distribution is ultimately asymmetric. A special case is that only the negation of a uniform probability distribution is uniform. At this time, this kind of unsupported information is the least, and a symmetrical distribution of negations is realized, as indicated by Theorem 4. On the other hand, when the negation parameter is the upper bound of negation, the reason for choosing the minimum entropy scheme is that it chooses one of the most accurate and intuitive alternatives. In this case, the uncertainty of the negation of the probability distribution is the smallest, which is more helpful for us to make decisions. This is similar to the state after the information is restored after defuzzification, and it can directly provide us with more reliable information.

Later, we shall use several numerical examples to further illustrate the characteristics of the proposed negation method, so as to impress the readers about the numerical meaning of getting the negation of a probability distribution.

3. Numerical examples

In this section, some numerical examples are provided to analyze the properties of the proposed negation procedure. In particular, Examples 1-3 are used to verify the theorems satisfied by the proposed scheme. In Example 4, we compare the proposed negation scheme with Yager's maximum entropy model and search some interesting findings.

EXAMPLE 1. *Assume a frame of reference is the set* $X = \{x_1, x_2, x_3, x_4, x_5\}$*. Let* P *be such that* $p_1 = 1$ *and* $p_i = 0$ *for* $i \neq 1$*. In this case not* P*, i.e.,* $\bar{P}$*, is obtained as shown in Table 1 and Figure 1.*

Table 1: The negation of the probability distribution in Example 1

$\bar{P}$	Negation parameter	
	$\kappa = 0$	$\kappa = 1$
$\bar{p}_1$	0.0225	0
$\bar{p}_2$	0.2444	0
$\bar{p}_3$	0.2444	0
$\bar{p}_4$	0.2444	0
$\bar{p}_5$	0.2444	1

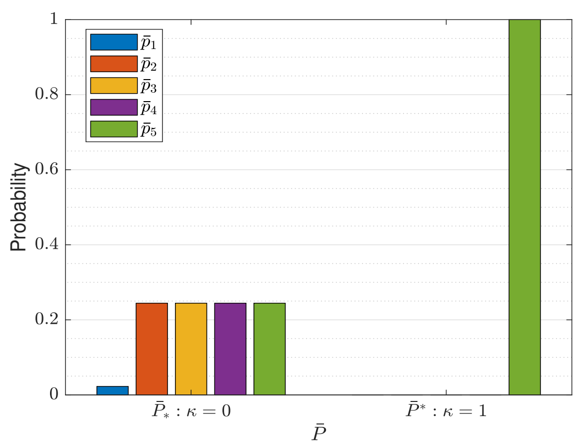

Figure 1: The graph on the negation of the probability distribution in Example 1.

In this example, the negation of the probability distribution P is provided at the lower and upper negation bounds, respectively. First of all, we can clearly observe that under these two different negation parameters, $\bar{P}_*$: $\sum_{i=1}^{5} \bar{p}_{i*} = 1$ with $\bar{p}_{i*} \in [0,1]$, as well as $\bar{P}^*$: $\sum_{i=1}^{5} \bar{p}_i^* = 1$ with $\bar{p}_i^* \in [0,1]$. That is, the negation of P is still a probability distribution, which verifies Theorem 1. In addition, we can see that $p_1 > p_i$ $(i \neq 1)$ in the initial probability distribution. After executing the negation procedure, in the lower and upper bounds of negation, the results we get are $\bar{p}_i \geq \bar{p}_1$ $(i \neq 1)$. In other words, if $p_i \geq p_j$ then $\bar{p}_i \leq \bar{p}_j$, which is in line with our intuition. Note that when the negation parameter is the lower bound of negation, Figure 1 only shows a possible probability allocation scheme. In fact, all elements except x_1 may get 1, since their initial probability distribution is the same. And this result depends on the position of the variable element x_i, which is determined by the index function. In addition, when the negation parameter is the lower bound of negation, we can observe that $H(\bar{P}_*)$ achieves a distribution close to the maximum entropy. When the negation parameter is the lower bound of negation, we can easily calculate $H(\bar{P}^*) = 0$. Moreover, we can get $H(\bar{P}_*) = 0.7606 > H(P) = 0$ through calculation. Therefore, this example verifies Theorems 1-2 and 4.

EXAMPLE 2. *Assume a frame of reference is the set $X = \{x_1, x_2, x_3, x_4, x_5\}$. Let P be such that $p_i = 1/5$. In this case not P, i.e., $\bar{P}$, is obtained as shown in Table 2 and Figure 2.*

Table 2: The negation of the probability distribution in Example 2

$\bar{P}$	Negation parameter	
	$\kappa = 0$	$\kappa = 1$
$\bar{p}_1$	0.2000	1
$\bar{p}_2$	0.2000	0
$\bar{p}_3$	0.2000	0
$\bar{p}_4$	0.2000	0
$\bar{p}_5$	0.2000	0

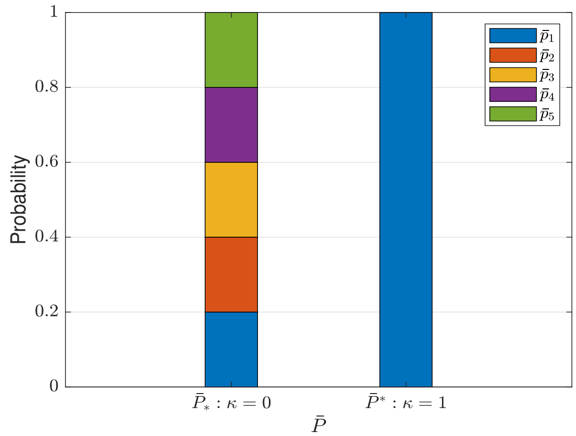

Figure 2: The graph on the negation of the probability distribution in Example 2.

In this example, what is given is a set of uniform probability distributions. We can find the following results. First of all, under these two different negation parameters, it is obvious that the negation of the original probability distribution is still a probability distribution, since $\sum_{i=1}^{5} \bar{p}_{i*} = 1$ with $\bar{p}_{i*} \in [0, 1]$, and $\bar{P}^*$: $\sum_{i=1}^{5} \bar{p}_i^* = 1$ with $\bar{p}_i^* \in [0, 1]$, which verify Theorem 1. In particular, at the lower bound of negation, it can be seen intuitively from Figure 2 that the negation of P is still uniformly distributed. That is the negation at this time achieves the maximum entropy, i.e., $H(\bar{P}_*)_{max}$. Thus, this also verifies Theorem 4. In addition, when the negation parameter is the upper bound of negation, a distribution way of the probability distribution is displayed as $\bar{p}_1 = 1$. This is similar to the case where the negation parameter is the upper bound of negation in Example 1, since the initial probability distribution of the five reference elements is the same. In this particular case, the negation of the probability distribution can be assigned to any element in the reference frame. In addition, by calculation we can obtain $H(\bar{P}_*) = H(P) = 0.8000$. Therefore, this example verifies Theorems 1-4, especially Theorem 3.

Example 3. *Assume a frame of reference is the set* $X = \{x_1, x_2, x_3, x_4, x_5\}$. *Let* $P = \{0.20, 0.30, 0.10, 0.25, 0.15\}$. *In this case not* P*, i.e.,* $\bar{P}$*, is obtained* *as shown in Table 3. In addition, Figure 2 shows the relationship between the negation of the probability distribution and entropy.*

Table 3: The negation of the probability distribution in Example 3

$\bar{P}$	Negation parameter κ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$\bar{p}_1$	0.2038	0.1956	0.1852	0.1726	0.1575	0.1396	0.1187	0.0946	0.0669	0.0354	0
$\bar{p}_2$	0.1766	0.1596	0.1424	0.1250	0.1074	0.0897	0.0718	0.0539	0.0359	0.0179	0
$\bar{p}_3$	0.2157	0.2524	0.2951	0.3448	0.4027	0.4698	0.5477	0.6378	0.7420	0.8620	1
$\bar{p}_4$	0.1918	0.1778	0.1627	0.1465	0.1291	0.1105	0.0907	0.0698	0.0477	0.0244	0
$\bar{p}_5$	0.2120	0.2146	0.2145	0.2110	0.2033	0.1904	0.1710	0.1440	0.1076	0.0603	0

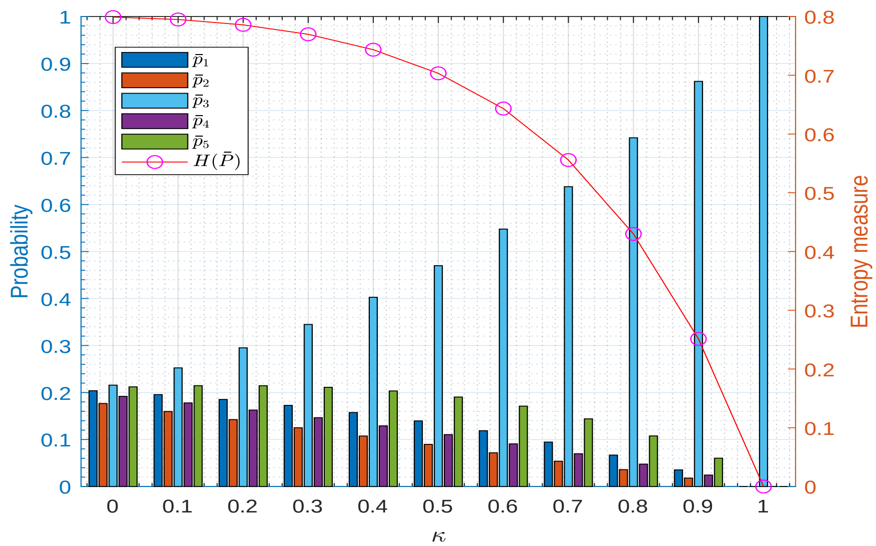

Figure 3: The graph on the negation of the probability distribution in Example 3.

In this example, we can see that when the strength of the negation is further deepened, that is, from the upper bound of the negation to the lower bound of the negation, the value of the element x_2 with the larger initial probability distribution is constantly decreasing, and the value of the element x_3 with the smaller initial probability distribution is always increase until $\bar{p}_3 = 1$. Moreover, in this process, we can observe that the value of entropy is decreasing after the negation of the probability distribution, until $H(\bar{P}^*) = 0$. In addition, we notice that under any of the above negation parameters, the probability distribution of the negation is still a probability distribution, because the constraints $\sum_{i=1}^{5} \bar{p}_i = 1$ and $\bar{p}_i \in [0,1]$ are always satisfied. In particular, in the lower bound of negation, we have $H(\bar{P}_*) = 0.7990 > H(P) = 0.7750$. Hence, this example verifies Theorems 1-2, and 4, especially Theorem 2.

Example 4. *Assume a frame of reference is the set $X = \{x_1, x_2\}$. Let P be such that $p_1 = 0.7$ and $p_2 = 0.3$.*

In this example, we let the initial probability distribution be the first iteration of negation. Then we calculate the probability distribution of different negation iterations under different negation parameters, and the results are shown in Table 4 and Figure 4.

Table 4: The change of the probability distribution (in Example 4) in the negation iteration under different negation parameters

κ	P	Frequency of negation												
		1	2	3	4	5	6	7	8	9	10	11	12	13
0	p_1	0.7	0.4013	0.5492	0.4754	0.5123	0.4939	0.5031	0.4985	0.5008	0.4996	0.5002	0.4999	0.5000
	p_2	0.3	0.5987	0.4508	0.5246	0.4877	0.5061	0.4969	0.5015	0.4992	0.5004	0.4998	0.5001	0.5000
0.1	p_1	0.7	0.3673	0.6009	0.4145	0.5781	0.4256	0.5726	0.4282	0.5713	0.4289	0.5710	0.4290	0.5710
	p_2	0.3	0.6327	0.3991	0.5855	0.4219	0.5744	0.4274	0.5718	0.4287	0.5711	0.4290	0.5710	0.4290
0.2	p_1	0.7	0.3319	0.6538	0.3527	0.6443	0.3570	0.6423	0.3579	0.6419	0.3581	0.6419	0.3581	0.6419
	p_2	0.3	0.6681	0.3462	0.6473	0.3557	0.6430	0.3577	0.6421	0.3581	0.6419	0.3581	0.6419	0.3581
0.3	p_1	0.7	0.2951	0.7070	0.2922	0.7081	0.2917	0.7083	0.2917	0.7083	0.2917	0.7083	0.2917	0.7083
	p_2	0.3	0.7049	0.2930	0.7078	0.2919	0.7083	0.2917	0.7083	0.2917	0.7083	0.2917	0.7083	0.2917
0.4	p_1	0.7	0.2569	0.7592	0.2349	0.7672	0.2321	0.7682	0.2317	0.7683	0.2317	0.7683	0.2317	0.2317
	p_2	0.3	0.7431	0.2408	0.7651	0.2328	0.7679	0.2318	0.7683	0.2317	0.7683	0.2317	0.7683	0.7683
0.5	p_1	0.7	0.2173	0.8095	0.1824	0.8200	0.1792	0.8210	0.1790	0.8210	0.1790	0.8210	0.1790	0.8210
	p_2	0.3	0.7827	0.1905	0.8176	0.1800	0.8208	0.1790	0.8210	0.1790	0.8210	0.1790	0.8210	0.1790
0.6	p_1	0.7	0.1764	0.8567	0.1354	0.8665	0.1331	0.8670	0.1330	0.8670	0.1330	0.8670	0.1330	0.8670
	p_2	0.3	0.8236	0.1433	0.8646	0.1335	0.8669	0.1330	0.8670	0.1330	0.8670	0.1330	0.8670	0.1330
0.7	p_1	0.7	0.1342	0.8999	0.0941	0.9069	0.0929	0.9071	0.0929	0.9071	0.0929	0.9071	0.0929	0.9071
	p_2	0.3	0.8236	0.1433	0.8646	0.1335	0.8669	0.1330	0.8670	0.1330	0.8670	0.1330	0.8670	0.1330
0.8	p_1	0.7	0.0906	0.9385	0.0583	0.9421	0.0579	0.9421	0.0579	0.9421	0.0579	0.9421	0.0579	0.9421
	p_2	0.3	0.9094	0.0615	0.9417	0.9421	0.9421	0.0579	0.9421	0.0579	0.9421	0.0579	0.9421	0.0579
0.9	p_1	0.7	0.0459	0.9719	0.0272	0.9729	0.0271	0.9729	0.0271	0.9729	0.0271	0.9729	0.0271	0.9729
	p_2	0.3	0.9541	0.0281	0.9728	0.0271	0.9729	0.0271	0.9729	0.0271	0.9729	0.0271	0.9729	0.0271
1	p_1	0.7	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
	p_2	0.3	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

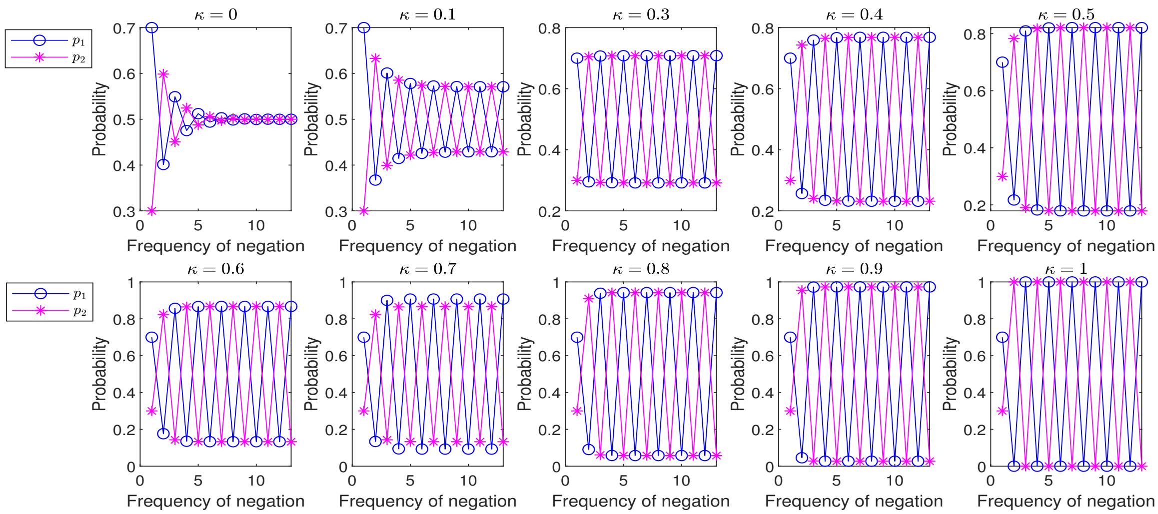

Figure 4: The graph on the negation of the probability distribution in Example 4 under different negation parameters.

We can observe an interesting discovery that under different negation parameters, with the iteration of negation, the probabilities of p_1 and p_2 are interchanged. In this case, the negation is reversible. In particular, when $\kappa = 0$, the maximum entropy distribution is realized at this time. In order to further find the reason for this phenomenon, we think from the perspective of entropy. We calculate the change in entropy of this probability distribution under different negation parameters, and the results are shown in Table 5 and Figure 5.

Table 5: The entropy change of the probability distribution (in Example 4) in the negation iteration under different negation parameters

κ	Frequency of negation											
	2	3	4	5	6	7	8	9	10	11	12	13
0	0.4805	0.4952	0.4988	0.4997	0.4999	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
0.1	0.4648	0.4796	0.4854	0.4878	0.4889	0.4894	0.4897	0.4898	0.4899	0.4899	0.4899	0.4899
0.2	0.4435	0.4527	0.4566	0.4583	0.4591	0.4595	0.4596	0.4597	0.4597	0.4597	0.4597	0.4597
0.3	0.4160	0.4143	0.4136	0.4134	0.4132	0.4132	0.4132	0.4132	0.4132	0.4132	0.4132	0.4132
0.4	0.3818	0.3656	0.3595	0.3573	0.3565	0.3562	0.3561	0.3561	0.3561	0.3561	0.3561	0.3561
0.5	0.3402	0.3084	0.2982	0.2951	0.2942	0.2940	0.2940	0.2940	0.2940	0.2940	0.2940	0.2940
0.6	0.2906	0.2455	0.2341	0.2314	0.2307	0.2306	0.2306	0.2306	0.2306	0.2306	0.2306	0.2306
0.7	0.2323	0.1801	0.1705	0.1688	0.1686	0.1686	0.1686	0.1686	0.1686	0.1686	0.1686	0.1686
0.8	0.1649	0.1155	0.1097	0.1091	0.1091	0.1091	0.1091	0.1091	0.1091	0.1091	0.1091	0.1091
0.9	0.0876	0.0547	0.0529	0.0528	0.0528	0.0528	0.0528	0.0528	0.0528	0.0528	0.0528	0.0528
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

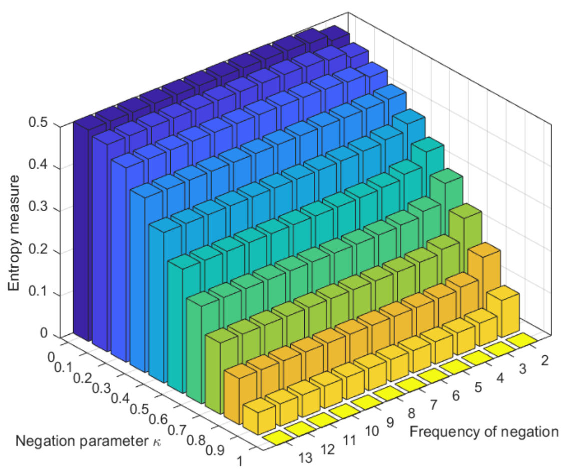

Figure 5: The relationship between entropy and negation parameter in the negation iteration process.

We can see that during the initial negation iteration, as the probability is distributed, the entropy of the negation probability distribution also changes. Until the probability of the element x_1 and x_2 is exchanged, the entropy reaches a certain value and does not change afterwards. This example shows that as the strength of the negation increases, the support for elements after each iteration of negation also increases. In essence, this phenomenon reflects the concept of mutual exchange of support before propositions. If we consider the maximum entropy model of Yager and the result is shown in Figure 6, we can see that in the 12 iterations of the negation, the probability distribution after negation only shows the exchange of probability, and $H(\bar{P})$ is always 0.42. In contrast, our negation method introduces the intensity level of negation execution based on the exchange probability. In contrast, our negation method introduces the intensity level of negation execution based on the exchange probability, so that the negation is constrained in a spatial domain, and the result is intuitive.

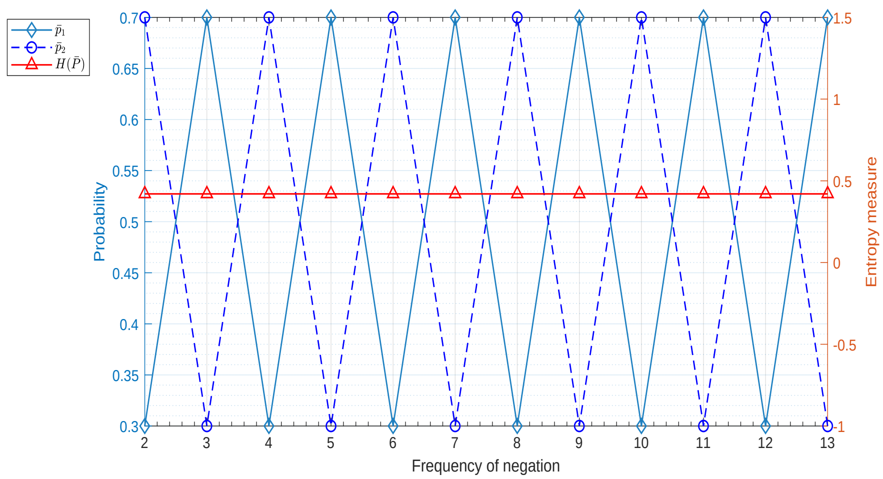

Figure 6: The negation calculation results of Yager's maximum entropy model to the probability distribution in Example 4.

4. View from Dempster-Shafer theory

Here, we shall further verify the actual connotation of the negation parameters in our proposed negation method from the perspective of Dempster-Shafer theory. But first, let us recall some basic concepts about this theory.

In Dempster-Shafer theory [20, 21], $X = \{x_1, x_2, \cdots, x_N\}$ represents the value space of a set of random variables, which is mutually exclusive and exhaustive, and is called a frame of discernment. For a given frame of discernment X, a belief structure m is denoted as a mapping, i.e., $m : 2^{\Omega} \to [0, 1]$, satisfying the following conditions:

$$m(\emptyset) = 0 \quad and \quad \sum_{F \subseteq X} m(F) = 1.$$

And if $m(F) > 0$, F is the focal element of the belief structure. $m(F)$ represents an assigned belief measure to accurately reflect the degree of support for F. The combination of all focus elements is called the core of the belief structure m. In addition, A *Bayesian* belief structure is such that all its focal elements are singletons elements of the frame of discernment [36, 37]. Multiple sources of evidence each represented by a belief structure can be combined by Dempster's combination rule. The rule to combine m_1, m_2, ..., m_N is defined as the following

$$[m_1 \oplus m_2 \cdots \oplus m_N](A) = \frac{\sum\limits_{\bigcap\limits_{i=1}^{N} F_i = A} \prod\limits_{i=1}^{N} m_i(F_i)}{1 - \sum\limits_{\bigcap\limits_{i=1}^{N} A_i = \emptyset} \prod\limits_{i=1}^{N} m_i(A_i)}$$

in which $A \neq \emptyset$, $F_i, A_i \in 2^{\Theta}$, whereas $m(\emptyset) = 0$ for $A = \emptyset$. One of the most important parameters in Dempster's combination rules is called conflict measure. For two given belief structures m_1 and m_2, the conflict between them is measured by

$$Conf = \sum_{F_i \cap F_j = \emptyset} m_1(F_i) m_2(F_j).$$

Generally, the larger $Conf$ is, the more conflict between these two belief structures. For instance, if $Conf = 0$, it indicates that there is no conflict between m_i and m_j. However, when $Conf = 1$, it indicates a complete conflict between m_1 and m_2.

Here we shall use the conflict coefficient $Conf$ from Dempster-Shafer theory to measure the degree of uncertainty between two belief structures. One reason for its choice is that since negation is intuitively back to the opposite side of this event, the negation model can at least satisfy the phenomenon that this event and its opposite event should be in complete conflict. Assume given such a *Bayesian* belief structure, expressed by

$$\mathbf{m} : m(\{F_1\}) = 0.9, m(\{T_2\}) = 0, m(\{T_3\}) = 0.1$$

its negation is denoted as $\bar{\mathbf{m}}$. Then we calculate the results of the proposed scheme and Yager's negation method. The data are shown in Table 6. The intuitive comparison of the two methods is shown in Figure 7.

Table 6: The negation of the belief structure $\mathbf{m}$

Method	Negation parameter	$Conf(\mathbf{m}, \bar{\mathbf{m}})$
Proposed method	$\kappa = 0$	0.8497
	$\kappa = 0.1$	0.8634
	$\kappa = 0.2$	0.8774
	$\kappa = 0.3$	0.8918
	$\kappa = 0.4$	0.9063
	$\kappa = 0.5$	0.9212
	$\kappa = 0.6$	0.9363
	$\kappa = 0.7$	0.9518
	$\kappa = 0.8$	0.9675
	$\kappa = 0.9$	0.9836
	$\kappa = 1$	1
Yager's method		0.9100

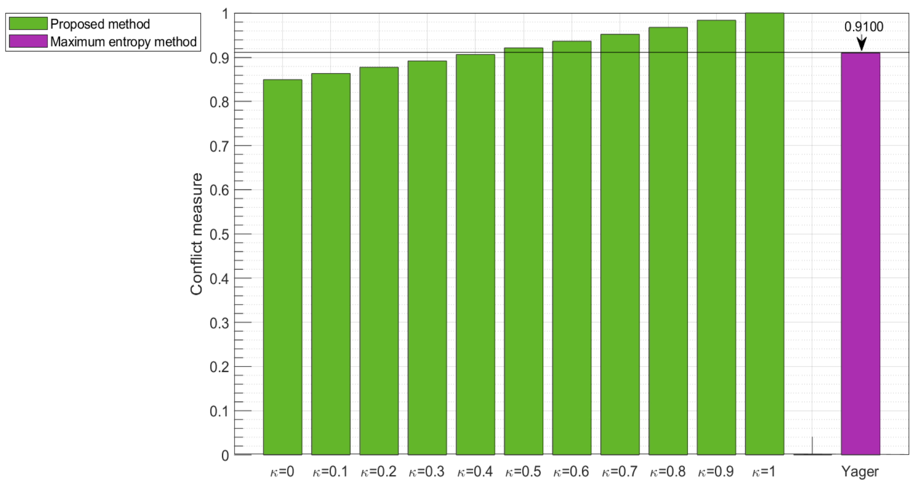

Figure 7: The graph on the negation of the belief structure $\mathbf{m}$.

Obviously, we can observe that using Yager's negation method, the conflict measure between belief structure $\mathbf{m}$ and $\bar{\mathbf{m}}$ is equal to 0.9100. In the proposed negation method, when $\kappa = 0.5$, the measure of conflict is greater than 0.9100. Moreover, as the level of negation increases later, when $\kappa = 1$, $Conf(\mathbf{m}, \bar{\mathbf{m}}) = 1$. In other words, when the negation parameter is the upper bound of negation, this initial belief structure $\mathbf{m}$ and its negation $\bar{\mathbf{m}}$ are completely in conflict.

Therefore, the proposed negation of a probability distribution is indeed a sound negation scheme, which considers all possible negations. This gives people a hope that through a change in the level of negation, the negation of a certain event I found is indeed the answer I want, and it is consistent with cognition and intuition.

5. Conclusions

In this paper, we advised a transformation method to obtain the negation of a probability distribution. Some negation operators are defined and some desired properties are proved. Numerical examples are used to demonstrate the unique features of the proposed method. We discovered the influence of the negation parameter in the negation space on the execution of the negation. Moreover, from the perspective of Dempster's belief structure, we verified the connotations of the negation parameter in the proposed negation scheme. Furthermore, we found that Yager's negation model seems to be one of the special cases of the negation scenario we proposal.

Acknowledgment

The Fund of the National Natural Science Foundation of China (Grant No. 61903307), Project funded by China Postdoctoral Science Foundation (Grant No. 2020M683575), and the Startup Fund from Northwest A&F University (Grant No. 2452018066), and the National College Students Innovation and Entrepreneurship Training Program (Grant No. S202010712135, No. S202010712019, No. X202010712364) are acknowledged for their financial support of this research.

Conflict of interest

The authors state that no conflicts of interest exist.

References

[1] L. A. Zadeh, Fuzzy sets, Information and control 8 (3) (1965) 338–353.

[2] Y. Xue, Y. Deng, H. Garg, Uncertain database retrieval with measure–based belief function attribute values under intuitionistic fuzzy set, Information Sciences 546 (2021) 436–447.

[3] R. Tao, Z. Liu, R. Cai, K. Cheong, A dynamic group mcdm model with intuitionistic fuzzy set: Perspective of alternative queuing method, Information Sciences 514 (2021) 85–103.

[4] Y. Deng, Information volume of mass function, International Journal of Computers, Communications & Control 15 (6) (2020) 3983.

[5] X. Mi, B. Kang, On the belief universal gravitation (bug), Computers & Industrial Engineering (2020) Doi: 10.1016/j.cie.2020.106685.

[6] F. Xiao, Ced: A distance for complex mass functions, IEEE Transactions on Neural Networks and Learning Systems (2020) Doi: 10.1109/TNNLS.2020.2984918.

[7] F. Xiao, Ceqd: A complex mass function to predict interference effects, IEEE Transactions on Cybernetics (2020) Doi: 10.1109/TCYB.2020.3040770.

[8] L. A. Zadeh, A note on z-numbers, Information Sciences 181 (14) (2011) 2923–2932.

[9] W. Jiang, Y. Cao, X. Deng, A novel z-network model based on bayesian network and z-number, IEEE Transactions on Fuzzy Systems 28 (8) (2020) 1585–1599.

[10] Q. Liu, Y. Tian, B. Kang, Derive knowledge of z-number from the perspective of dempster-shafer evidence theory, Engineering Applications of Artificial Intelligence 85 (2019) 754–764.

[11] J. Yang, D. Xu, Evidential reasoning rule for evidence combination, Artificial Intelligence 205 (2013) 1–29.

[12] L. Fei, Y. Deng, Multi-criteria decision making in pythagorean fuzzy environment, Applied Intelligence 50 (2) (2020) 537–561.

[13] X. Mi, Y. Tian, B. Kang, A hybrid multi-criteria decision making approach for assessing health-care waste management technologies based on soft likelihood function and d-numbers, Applied Intelligence (2020) Doi: 10.1007/s10489–020–02148–7.

[14] F. Xiao, Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy, Information Fusion 46 (2019) 23–32.

[15] F. Xiao, A new divergence measure for belief functions in d–s evidence theory for multisensor data fusion, Information Sciences 514 (2019) 462–483.

[16] X. Mi, T. Lv, Y. Tian, B. Kang, Multi-sensor data fusion based on soft likelihood functions and owa aggregation and its application in target recognition system, ISA transactions (2020) Doi: 10.1016/j.isatra.2020.12.009.

[17] R. R. Yager, L. A. Zadeh, An introduction to fuzzy logic applications in intelligent systems, Vol. 165, Springer Science & Business Media, 2012.

[18] Y. Song, J. Zhu, L. Lei, X. Wang, A self-adaptive combination method for temporal evidence based on negotiation strategy, SCIENCE CHINA Information Sciences 63 (11) (2020) 56–68.

[19] C. Baudrit, D. Dubois, D. Guyonnet, Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment, IEEE transactions on fuzzy systems 14 (5) (2006) 593–608.

[20] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (2) (1967) 325–339.

[21] G. Shafer, Dempster-shafer theory, Encyclopedia of artificial intelligence 1 (1992) 330–331.

[22] L. A. Zadeh, A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination, AI magazine 7 (2) (1986) 85–85.

[23] R. R. Yager, On the maximum entropy negation of a probability distribution, IEEE Transactions on Fuzzy Systems 23 (5) (2014) 1899–1902.

[24] L. Yin, X. Deng, Y. Deng, The negation of a basic probability assignment, IEEE Transactions on Fuzzy Systems 27 (1) (2018) 135–143.

[25] Z. Luo, Y. Deng, A matrix method of basic belief assignment's negation in dempster-shafer theory, IEEE Transactions on Fuzzy Systems 28 (9) (2019) 2270–2276.

[26] X. Deng, W. Jiang, On the negation of a dempster–shafer belief structure based on maximum uncertainty allocation, Information Sciences 516 (2020) 346–352.

[27] D. Wu, Z. Liu, Y. Tang, A new classification method based on the negation of a basic probability assignment in the evidence theory, Engineering Applications of Artificial Intelligence 96 (2020) 103985.

[28] S. Li, F. Xiao, J. H. Abawajy, Conflict management of evidence theory based on belief entropy and negation, IEEE Access 8 (2020) 37766–37774.

[29] Q. Liu, H. Cui, Y. Tian, B. Kang, On the negation of discrete z-numbers, Information Sciences (2020) Doi: 10.1016/j.ins.2020.05.106.

[30] A. Srivastava, S. Maheshwari, Some new properties of negation of a probability distribution, International Journal of Intelligent Systems 33 (6) (2018) 1133–1145.

[31] J. Zhang, R. Liu, J. Zhang, B. Kang, Extension of yager's negation of a probability distribution based on tsallis entropy, International Journal of Intelligent Systems 35 (1) (2020) 72–84.

[32] X. Gao, Y. Deng, The generalization negation of probability distribution and its application in target recognition based on sensor fusion, International Journal of Distributed Sensor Networks 15 (5) (2019) 1550147719849381.

[33] A. Srivastava, L. Kaur, Uncertainty and negation—information theoretic applications, International Journal of Intelligent Systems 34 (6) (2019) 1248–1260.

[34] F. Xiao, On the maximum entropy negation of a complex-valued distribution, IEEE Transactions on Fuzzy Systems (2020) Doi: 10.1109/TFUZZ.2020.3016723.

[35] C. E. Shannon, A mathematical theory of communication, The Bell system technical journal 27 (3) (1948) 379–423.

[36] P. Smets, Decision making in the tbm: the necessity of the pignistic transformation, International Journal of Approximate Reasoning 38 (2) (2005) 133–147.

[37] W. Jiang, C. Huang, X. Deng, A new probability transformation method based on a correlation coefficient of belief functions, International Journal of Intelligent Systems 34 (6) (2019) 1337–1347.