Discounting or Optimizing? Different Approaches to Pseudo-Belief Function Correction

Milan Daniel[1], Radim Jiroušek[2], and Václav Kratochvíl[2]

[1]Institute of Computer Sciences, Czech Academy of Sciences
[1]*milan.daniel@cs.cas.cz*
[2]Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic
[2] {*radim;velorex*}*@utia.cas.cz*

Abstract

We present and compare several approaches for transforming pseudo-belief functions, constructed from Jeffreys confidence intervals on observational data, into proper belief functions. Two main classes of methods are examined: one based on polyhedral geometry using various optimization strategies, and the other employing generalized belief discounting. Finally, the proposed methods are evaluated on real cybersecurity data and compared with standard upper and lower approximations of pseudo-belief.

1 Introduction

In a companion paper in these proceedings Daniel et al. (2025), belief functions were estimated from data using lower bounds based on Jeffrey's binomial confidence intervals. These bounds may not directly correspond to any valid belief function, leading to the notion of *pseudo-belief functions* — representations that preserve the intended epistemic meaning but violate some mathematical constraints of belief calculus.

This paper presents two complementary groups of methods for correcting such pseudo-belief functions and obtaining valid belief functions from them. Each method has its own motivation and interpretation:

- The first approach is based on *polyhedral geometry*. It considers the lower bounds (e.g., Jeffreys-type) as defining a polyhedron of admissible belief functions and selects one or more representative elements from this set using geometric or optimization-based criteria. This approach is constructive and data-driven.

- The second approach, introduced in this paper, generalizes the classical method of *belief discounting* as defined by Shafer (1976). It assumes a partial reliability of the original pseudo-belief function and proportionally reduces its support, originally transferring the remainder to total ignorance. We utilize the remainder for negative belief mass correction here. The result is a valid belief function that retains the internal structure of the original function.

While these two approaches differ in spirit — geometric reconstruction versus numerical correction — they share the same objective: to obtain valid belief functions that are compatible with uncertain or incomplete evidence. Moreover, belief discounting is a linear transformation and can be interpreted as a special case of movement within the credal set (i.e., the set of all belief functions consistent with given information), hence admitting a geometric interpretation.

By combining these perspectives, the paper contributes to the broader effort of belief function learning: deriving reasonable epistemic representations from empirical data, even in the presence of imprecision or ambiguity.

2 Preliminaries

This paper builds on a preceding contribution in these proceedings (Daniel et al., 2025), where belief and pseudo-belief functions were introduced and motivated by data-driven lower bounds such as Jeffreys intervals. These bounds may not always define a valid belief function, leading to pseudo-belief structures that require correction.

We explore two complementary correction strategies, each developed in a separate section. The first is based on polyhedral geometry and is presented next. The second relies on belief discounting and follows afterward. Here we briefly recall the key concepts common to both.

A belief function over a finite frame Ω is defined via a basic probability assignment (BPA) $m : 2^\Omega \to [0, 1]$ satisfying $m(\emptyset) = 0$, $\sum m(A) = 1$. Belief and plausibility functions are given by:

$$\mathrm{bel}_m(A) = \sum_{B \subseteq A} m(B), \quad \mathrm{pl}_m(A) = \sum_{B \cap A \neq \emptyset} m(B). \tag{1}$$

Pseudo-belief functions generalize this by allowing some negative masses while preserving the belief–plausibility duality.

An important correction tool is *discounting* (Shafer, 1976), used when evidence is not fully reliable. Given trust $1 - \delta$, the discounted mass function is defined as:

$$m^\delta(A) = (1 - \delta)m(A) \text{ for } A \neq \Omega, \quad m^\delta(\Omega) = (1 - \delta)m(\Omega) + \delta,$$

yielding a belief function bel^δ with $\mathrm{bel}^\delta(A) = (1 - \delta)\,\mathrm{bel}(A)$ for $A \neq \Omega$, $\mathrm{bel}^\delta(\Omega) = 1$. Discounting increases uncertainty while preserving belief ratios.

Finally, when belief or plausibility bounds are defined by inequalities, they form a polyhedron

$$\mathcal{P} = \{x \in \mathbb{R}^n : Mx \leq b\},$$

which may or may not correspond to any belief function. We use geometric methods to modify such sets into valid belief structures — a topic of the next section.

3 Geometric Correction via Polyhedral Optimization

In the preceding chapters, we introduced pseudo-belief functions derived from empirical data using Jeffreys confidence intervals, as motivated in Daniel et al. (2025). These define lower and upper bounds for the belief and plausibility of each subset $A \subseteq \Omega$, resulting in a system of linear inequalities that constrains possible belief functions.

Let $\mathbf{bel}_J \in \mathbb{R}^{2^n}$ denote the vector of lower bounds (Jeffreys intervals) for each subset A. The inequality

$$\mathrm{bel}_J(A) \leq \sum_{B \subseteq A} m(B)$$

can be rewritten in matrix form as

$$Mm \geq \mathbf{bel}_J,$$

where $m \in \mathbb{R}^{2^n}$ is a vector of bpa values and $M \in \{0,1\}^{2^n \times 2^n}$ is the inclusion matrix with entries $M_{[A,B]} = 1$ iff $B \subseteq A$.

If we add constraints for normalization and non-negativity,

$$\sum_{A \subseteq \Omega} m(A) = 1, \qquad m(A) \geq 0,$$

we obtain a polytope $\mathcal{P}_J^* \subset \mathbb{R}^{2^n}$ containing all belief functions consistent with the empirical bounds. The polytope may have many vertices, corresponding to different consistent BPAs.

Polyhedral geometry offers strong tools for selecting one specific point $m \in \mathcal{P}_J^*$ by optimizing a suitable objective function. This approach draws on the convex geometry of belief spaces as explored in Cuzzolin (2010, 2020) and relies on standard polyhedral optimization methods (Ziegler, 1995; Bagnara et al., 2008).

We consider four optimization criteria:

Zero Objective (ZO). Selects any feasible point:

$$\text{Minimize } f_{\mathrm{ZO}}(m) = 0.$$

Sparsity (SP). Minimizes the number of focal elements:

$$\text{Minimize } f_{\mathrm{SP}}(m) = \sum_{A \subseteq \Omega, A \neq \emptyset} \delta[m(A) > 0],$$

where $\delta[\cdot]$ is the indicator function. This is approximated in LP by introducing binary variables z_A and constraints $m(A) \leq M \cdot z_A$, for large M.

Cardinality-Weighted (CW). Penalizes small subsets:

$$\text{Minimize } f_{\mathrm{CW}}(m) = \sum_{A \subseteq \Omega, A \neq \emptyset} \frac{1}{|A|} m(A).$$

Dubois–Prade Entropy (HD). Maximizes entropy:

$$\text{Maximize } H_D(m) = \sum_{A \subseteq \Omega, A \neq \emptyset} m(A) \log |A|,$$

as proposed in Dubois and Prade (1987).

The resulting belief mass assignments under each objective are shown below for the example with $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$:

Table 1: Comparison of belief mass assignments under different objective functions. Column m_J shows the pseudo-belief mass vector constructed directly from Jeffreys intervals (Daniel et al., 2025).

A	$\text{bel}_J(A)$	$m_J(A)$	$m_{ZO}(A)$	$m_{SP}(A)$	$m_{CW}(A)$	$m_{HD}(A)$
$\{\omega_1\}$	0.164	0.163	0.400	0.163	0.164	0.174
$\{\omega_2\}$	0.211	0.211	0.321	0.297	0.211	0.222
$\{\omega_3\}$	0.179	0.179	0.229	0.317	0.179	0.190
$\{\omega_4\}$	0.050	0.050	0.050	0.222	0.050	0.060
$\{\omega_1, \omega_2\}$	0.461	0.086	0	0	0.081	0.065
$\{\omega_1, \omega_3\}$	0.423	0.080	0	0	0.075	0.059
$\{\omega_1, \omega_4\}$	0.262	0.048	0	0	0.044	0.027
$\{\omega_2, \omega_3\}$	0.480	0.089	0	0	0.089	0.068
$\{\omega_2, \omega_4\}$	0.314	0.052	0	0	0.052	0.031
$\{\omega_3, \omega_4\}$	0.279	0.050	0	0	0.050	0.029
$\{\omega_1, \omega_2, \omega_3\}$	0.778	-0.031	0	0	0	0
$\{\omega_1, \omega_2, \omega_4\}$	0.580	-0.032	0	0	0	0
$\{\omega_1, \omega_3, \omega_4\}$	0.539	-0.032	0	0	0	0
$\{\omega_2, \omega_3, \omega_4\}$	0.600	-0.032	0	0	0	0
Ω	1.000	0.117	0	0	0	0.075
H_D			0	0	0.271	0.297

Discussion. Each optimization criterion leads to a different belief assignment. The **Sparsity (SP)** solution concentrates belief on a minimal number of focal elements, which may be beneficial for interpretability. The **Cardinality-Weighted (CW)** solution favors larger sets, thereby reflecting cautious reasoning. The **Dubois–Prade entropy (HD)** solution maximizes epistemic uncertainty and spreads mass over larger focal elements, constrained by empirical bounds. Interestingly, the **CW** solution closely resembles the approximation obtained in (Daniel et al., 2025) using the Upper Approximation Procedure, which supports the validity of our optimization-based approach.

These results illustrate that geometric correction methods not only ensure consistency with empirical estimates but also allow tailoring belief functions according to different modeling principles or user preferences.

4 A Generalization of Belief Discounting

Let us assume a PBF bel_J constructed by application of the method of Jeffreys confidence intervals (Daniel et al., 2025) or any general PBF, which does not satisfy the classic Shafer's definition of BF - there are some focal elements in respective BPA with negative mass, cf. negative ε in lower approximation in Daniel et al. (2025). Our aim is simply to find an acceptable way how to eliminate these negative belief masses. From the definition of a PBF, all pseudo-beliefs are non-negative, thus also all belief masses of singletons are obviously non-negative. Hence, an issue can appear only for $X \subseteq \Omega$, $|X| \geq 2$.

We can consider the negative mass $m(X)$ as a consequence of over information about the case which is the subject of the belief; excess information on subsets of X, which can be corrected by removal of at least belief mass corresponding to the sum of negative belief masses.

Example 1 In the simplest case of such a PBF, i.e., bel_S on $|\Omega_2| = 2$, $m_S(\Omega_2) = \varepsilon < 0$, we can simply solve the problem by discounting with discounting rate $\delta \geq -\varepsilon/(1-\varepsilon)$: $m_S^\delta(\Omega_2) = (1-\delta)m_S(\Omega_2) + \delta \geq (1-(-\varepsilon/(1-\varepsilon))\varepsilon - \varepsilon/(1-\varepsilon) = \varepsilon(1+\varepsilon/(1-\varepsilon) - 1/(1-\varepsilon) = \varepsilon(1-1) = 0$, $m_S^\delta(\omega_i) = (1-\delta)m_S(\omega_i)$, which is also non-negative, as $1-\delta = 1+\varepsilon/(1-\varepsilon) = (1-\varepsilon)/(1-\varepsilon) + \varepsilon/(1-\varepsilon) = 1/(1-\varepsilon) > 0$. Just the same procedure we can use for correction of any PBF on $|\Omega_n| = n$, with the only negative belief mass $m(\Omega_n)$.

In general, we have to correct all negative belief masses, of any focal element $|X| > 1$. We can do it in three different ways:

(i) *local correction* of $m(X)$, related to the only focal element $X \subset \Omega$ and its subsets,
(ii) *layered correction* of all $m(X)$, s.t. $|X| = k$, related only to focal elements $|Y| \leq k$,
(iii) *global correction* which correct all the focal elements with negative masses together.

Motivated by the above solution of the simplest PBF case we will try to generalize discounting as it follows.

Local discounting on $X \subset \Omega$: $m^{\ulcorner X \urcorner \delta}(A) = (1-\delta)m(A)$ for any $A \subset X$, $m^{\ulcorner X \urcorner \delta}(X) = m(X) + \delta \sum_{A \subset X} m(A)$ and $m^{\ulcorner X \urcorner \delta}(A) = m(A)$ for other subsets $A \subseteq \Omega$, i.e., for $A \not\subseteq X$.
Property 1: If X is disjoint with all focal elements of the same and less cardinality which are not its subsets ($X \cap Y = \emptyset$ for all $|Y| \leq |X|$ s.t. $Y \not\subseteq X$) we can describe reasonable property of this definition of local discounting: ($bel^{\ulcorner X \urcorner \delta}(A) = (1-\delta)bel(A)$ for any $A \subset X$, $bel^{\ulcorner X \urcorner \delta}(A) = bel(A)$ if A $\not\subset X$).
Property 2: If X is disjoint with all focal elements of the same cardinality, we can describe the following property of this definition of local discounting: ($bel^{\ulcorner X \urcorner \delta}(A) = (1-\delta)bel(A)$ for any $A \subset X$, $(1-\delta)bel(A) \leq bel^{\ulcorner X \urcorner \delta}(A) \leq bel(A)$ for any $|A| < |X| : A \not\subset X \& A \cap X \neq \emptyset$, $m^{\ulcorner X \urcorner \delta}(A) = m(A)$ otherwise.

Example 2 Let suppose $|\Omega_7| = 7$, with only two focal elements of cardinality 3, $|A_i| = 3$: $A_1 = \{\omega_1, \omega_2, \omega_3\}$, $A_2 = \{\omega_4, \omega_5, \omega_6\}$. Property 2 holds for both $bel^{\ulcorner A_i \urcorner \delta}$. If further $m(X) = 0$ for any $X \subseteq A_1 \cap A_2$, property 1 also holds for both $bel^{\ulcorner A_i \urcorner \delta}$. If there are added f.e.s $A_3 = \{\omega_4, \omega_5, \omega_7\}$, $A_4 = \{\omega_5, \omega_6, \omega_7\}$, the properties does hold only for $bel^{\ulcorner A_2 \urcorner \delta}$.

Hence, the above useful properties does not hold for general PBFs. Jeffreys bel_J has often focal elements intersecting with the others of the same cardinality. Thus for our reason the definition of local discounting should be improved in the future.

Cardinality or k-discounting for $1 < k \leq n$: $m^{k\delta}(A) = (1-\delta)m(A)$ for any $A : |A| < k$, $m^{k\delta}(A) = m(A) + \delta \sum_{B \subset A} \frac{m(A)}{\sum_{B \subset C : |C| = k} m(C)} m(B)$ for any $|A| = k$, and $m^{k\delta}(A) = m(A)$ for any $|A| > k$.

The formula - more precisely the coefficient of $m(B)$ seems to be rather complicated here, nevertheless we need to distribute any $m(B)$ among subsets of cardinality k, resp. just among all C, such that $B \subset C$ & $|C| = k$.

Observation 1 We can observe that $bel^{k\delta}(A) = (1-\delta)bel(A)$ for $|X| < k$ and $(1-\delta)bel(A) \leq bel^{k\delta}(A) \leq bel(A)$ for $|A| \geq k$: the first equality holds for $m(A) = 0$ and the second if there is the only one focal element of cardinality k, specially for $k = n$. [To be proved]

Observation 2 We can observe $m^{\lceil \Omega_n \rceil \delta}(A) = bel^{n\delta}(A) = bel^{\delta}(A)$ for $|\Omega_n| = n$. Thus both local and cardinality discounting are generalization of the original Shafer's discounting.

Proof. $m^{\lceil \Omega_n \rceil \delta}(\Omega_n) = m(\Omega_n) + \delta \cdot \sum_{A \subset \Omega_n} m(A) = (1-\delta)m(\Omega_n) + \delta m(\Omega_n) + \delta \cdot \sum_{A \subset \Omega_n} m(A) = (1-\delta)m(\Omega_n) + \delta(m(\Omega_n) + \sum_{A \subset \Omega_n} m(A)) = (1-\delta)m(\Omega_n) + \delta = m^{\delta}(\Omega_n)$. $m^{n\delta}(\Omega_n) = m(\Omega_n) + \delta \cdot \sum_{B \subset \Omega_n} m(\Omega_n)/m(\Omega_n) \cdot m(B) = (1-\delta)m(\Omega_n) + \delta m(\Omega_n) + \delta \cdot \sum_{B \subset \Omega_n} m(B)) = (1-\delta)m(\Omega_n) + \delta = m^{\delta}(\Omega_n)$.

5 Reduction of Over-Belief by Generalized Discounting

5.1 General Remarks on PBF Correction

Motivated by the successful Example 1, we have generalized belief discounting to allow analogous correction of general PBFs in the previous section.

As we have not yet been fully successful with the generalization of local discounting for general PBFs while preserving the ratios of belief masses, we resort to using only cardinality-based discounting in our corrections here. It affects all focal elements of a given cardinality (it distributes the discounted belief mass among all of them) and, similarly to classical Shafer discounting, adds the discounted mass to only one cardinality. Thus, our local and global corrections are in fact mixtures of local/global and layered corrections.

We must correct all belief masses < 0, i.e., such that $\sum_{B \subset A} m(B) > bel(A)$, i.e., with ratio $R(A) = bel(A) / \sum_{B \subset A} < 1$. If the lowest ratio among focal elements of cardinality k is used, then the strongest correction is performed, and the entire cardinality level is corrected. If the highest ratio is used, the smallest correction is performed, and only the focal element with that ratio is corrected. Note that a more complex formula for distributing the discounted belief mass is applied in cardinality discounting compared to classical discounting, and thus determining δ is also more complex, even though the underlying idea is analogous to that in Example 1.

5.2 Local, Layered, and Global Corrections of PBFs

1 Local correction should be the most precise, nevertheless it appears more complicated both from the theoretical point of view and also due to its computational complexity. As the theoretical part is not yet fully investigated, we adopt a mixture of local and layered approaches, correcting entire cardinality levels or individual focal elements one by one.

2 Layered correction is a compromise approach that corrects entire cardinality levels.

3 Global correction should correct all negative belief masses of the entire PBF together, if possible. Nevertheless, we again use only a mixture with layered correction.

5.3 Local Correction Algorithms

5.3.1 Algorithm 1-ugr

For each cardinality with negative pseudo-belief mass(es), we repeatedly utilize cardinality discounting with minimal correction discount rates (i.e., upward from minimal correction), correcting focal elements of cardinality k one by one using discount rates $\delta_A = -\sum_{|B|=k} m(B)/\sum_{C\subset A} m(C)$ for $|A| = k$, ordered from minimal to maximal, correcting pseudo-belief mass $m(A)$ if it is negative.

Algorithm 1

Compute pseudo m_J by Möbius transformation from bel_J, i.e., from the lower bounds of estimated confidence intervals;
For all $X \subseteq \Omega_n$:
$m_1(X) := m_J(X)$,
$R(X) := \min(1, bel_J(X)/\sum_{Y\subset X} m_J(X))$ for $|X| > 1$, $bel_J(X) > 0$,
$R(X) := 1$ for $|X| = 1$ or $bel_J(X) = 0$ or $\sum_{Y\subset X} m_J(X) = 0$.
$n := |\Omega_n|$;
For $k = 2, ..., n$:
$r_k := \max_{|X|=k} R(X)$,
If $r_k < 1$ **Then** $RFE := \{X \subseteq \Omega_n \mid |X| = k\}$
While $RFE \neq \emptyset$:
$AFE := \{X \in RFE \mid \text{with max } \sum_{Y\in RFE} m_{k-1}(Y)\}$
$sum_A := \sum_{B\subset A} m_{k-1}(B)$ for some $A \in AFE$,
$\delta_A := -\sum_{|B|=k} m_{k-1}(B)/sum_A$,
$m_k(X) := (1 - \delta_A) \cdot m_{k-1}(X)$ for $|X| < k$,
$m_k(X) := m_{k-1}(X) + \left|m_{k-1}(X) \cdot \sum_{Y\subset X} m(Y)/sum_A\right|$ for $|X| = k$,
$m_k(\Omega_n) := m(\Omega_n) + \sum_{|X|=k} \left(m_{k-1}(X)/sum_A \cdot \sum_{Y\not\subset X} m_{k-1}(Y)\right)$
Else For all $X \subseteq \Omega_n$: $m_k(X) := m_{k-1}(X)$;
% Now, $m(X) \geq 0$ for all X s.t. $|X| \leq k$
For all $X \subseteq \Omega_n$: $m(X) := m_n(X)$.

In the case of our cybersecurity data example from Daniel et al. (2025), there are four negative pseudo-belief masses of 3-element focal elements: $m(\{\omega_1, \omega_2, \omega_3\}) = -0.03139$, $m(\{\omega_1, \omega_2, \omega_4\}) = -0.03158$, $m(\{\omega_1, \omega_3, \omega_4\}) = -0.03159$, and $m(\{\omega_2, \omega_3, \omega_4\}) = -0.03157$ (see Table 1 and also the red values in Table 2). The corresponding discount rates are: $\delta_{(123)} = 0.155935$, $\delta_{(234)} = 0.044932$, $\delta_{(124)} = 0.002343$, $\delta_{(134)} = 0.000130$. For the final mass assignment further denoted as m_1 and the corresponding BF bel_1, see Tables 2 and 3. The belief function bel_1 is in fact a composition of four 3-discountings:

$$bel_1 = (((bel_J^{3\delta_{(123)}})^{3\delta_{(234)}})^{3\delta_{(124)}})^{3\delta_{(134)}}.$$

5.3.2 Algorithm 1-dgr

We utilize cardinality discounting with the maximal correction rate (i.e., the highest rate necessary to correct all negative belief masses of focal elements of cardinality k). Since this correction always affects all focal elements of cardinality k simultaneously, it essentially corresponds to Algorithm 2 from the next subsection.

5.3.3 Algorithms 1-ulr and 1-dlr

These algorithms aim to be closer to truly local corrections, either upward from the minimal or downward from the maximal correction. However, it is still under investigation whether it is possible to define an improved version of local discounting that preserves belief mass proportions as much as possible.

5.4 Layered Correction Algorithm

We apply cardinality discounting with the maximal discount rate, i.e., the minimal rate that corrects all negative belief masses of focal elements of a given cardinality k. This correction thus always affects all focal elements of that cardinality together.

In the case of our cybersecurity data example from Daniel et al. (2025), the maximal discount rate is $\delta_{(134)} = 0.2211$. Since negative pseudo-belief masses only appear at cardinality 3, the resulting BF further denoted bel_2 is a simple application of 3-discounting: $bel_2 = bel_J^{3\delta_{(134)}}$, see Tables 2 and 3. Note that this discount rate corresponds to $\{\omega_1, \omega_3, \omega_4\}$, which was the last focal element corrected in Algorithm 1-ugr. Nevertheless, the discount rate differs because here it is applied directly to the original bel_J, whereas in Algorithm 1-ugr it was applied after three previous corrections.

Algorithm 2. Layered Correction

Compute pseudo m_J by Möbius transformation from bel_J;
For all $X \subseteq \Omega_n$:
$m_1(X) := m_J(X)$,
$R(X) := \min(1, bel_J(X)/\sum_{Y \subset X} m_J(X))$ for $|X| > 1$, $bel_J(X) > 0$,
$R(X) := 1$ for $|X| = 1$ or $bel_J(X) = 0$ or $\sum_{Y \subset X} m_J(X) = 0$,
$n := |\Omega_n|$;
For $k = 2, \ldots, n$:
$r_k := \min_{|X|=k} R(X)$,
If $r_k < 1$ **Then**
$sum_\delta := \min_{|X|=k} \sum_{B \subset X} m_{k-1}(B)$,
$\delta_k := -\sum_{|B|=k} m_{k-1}(B)/sum_\delta$,
$m_k(X) := (1 - \delta_k) \cdot m_{k-1}(X)$ for $|X| < k$,
$m_k(X) := m_{k-1}(X) + |m_{k-1}(X) \cdot \sum_{Y \subset X} m_{k-1}(Y)/sum_\delta|$ for $|X| = k$,
$m_k(\Omega) := m_{k-1}(\Omega_n) + \sum_{|X|=k} \Big(m_{k-1}(X)/sum_\delta \cdot \sum_{Y \not\subset X} m_{k-1}(Y) \Big)$,
Else For all $X \subseteq \Omega_n$: $m_k(X) := m_{k-1}(X)$;
% $m_{k-1}(X) \geq 0$ for all $|X| \leq k$
For all $X \subseteq \Omega_n$: $m(X) := m_n(X)$.

5.5 Global Correction Algorithm(s)

Unfortunately, we do not yet have a truly global correction. Since cardinality discounting only corrects one cardinality level at a time, this method corresponds to an upside-down layered discounting — starting from cardinality n downward.

In the cybersecurity data example, there is only one cardinality (3) with negative pseudo-belief masses. Thus, the result of this approach is again $bel_3 = bel_J^{3\delta_{(134)}}$.

Algorithm 3. Global Correction

Compute pseudo m_J by Möbius transformation from bel_J; $n := |\Omega_n|$;
For all $X \subseteq \Omega_n$:
$m_1(X) := m_J(X)$,
$R(X) := \min(1, bel_J(X)/\sum_{Y \subset X} m_J(X))$ for $|X| > 1$, $bel_J(X) > 0$,
$R(X) := 1$ for $|X| = 1$ or $bel_J(X) = 0$ or $\sum_{Y \subset X} m_J(X) = 0$;
For $k = n, n-1, \ldots, 2$:
$r_k := \min_{|X|=k} R(X)$,
If $r_k < 1$ **Then**
$sum_\delta := \min_{|X|=k} \sum_{B \subset X} m_{k-1}(B)$,
$\delta_k := -\sum_{|B|=k} m_{k-1}(B)/sum_\delta$,
$m_k(X) := (1 - \delta_k) \cdot m_{k-1}(X)$ for $|X| < k$,
$m_k(X) := m_{k-1}(X) + |m_{k-1}(X) \cdot \sum_{Y \subset X} m_{k-1}(Y)/sum_\delta|$ for $|X| = k$,
$m_k(\Omega) := m_{k-1}(\Omega_n) + \sum_{|X|=k} \Big(m_{k-1}(X)/sum_\delta \cdot \sum_{Y \not\subset X} m_{k-1}(Y) \Big)$,
Else For all $X \subseteq \Omega_n$: $m_k(X) := m_{k-1}(X)$;
% $m_{k-1}(X) \geq 0$ for all $|X| \leq k$
For all $X \subseteq \Omega_n$: $m(X) := m_n(X)$.

5.5.1 Alternative Algorithms

There are two ideas for alternative algorithms. The first is a layered approach that mirrors Algorithm 1 in reverse: stepwise correction from the smallest to the largest $R(X)$ within each cardinality, moving downward from n to 2. The second is an open question: whether it is possible to define some generalized discounting operation that can correct all "negative" cardinalities at once.

6 Comparison on Cybersecurity Data Example

Let us compare lower and advanced lower approximations described in Daniel et al. (2025) with the approaches studied here. Specially, with geometric cardinality-weighted minimization (CW) and Dubois-Prade entropy (HD) and also with local and layered correction based on generalized discounting.

As all studied approaches are just in the process of their development, also our implementations are still in progress. Thus, we currently have correct comparable results only on 4-element frames of discernment now. General procedures are still in the middle of their tuning. Hence we will compare our approaches on the simplest case defined on cybersecurity data by Table 2 in Daniel et al. (2025): having 52 data records on the

4-element frame of discernment. For approximated/corrected belief mass assignments see Table 2, for approximated/corrected BFs see Table 3. Note that indices of m_1, m_2, bel_1, bel_2 refer results of Algorithms 1 and 2 here, not intermediate steps of their processing.

We skip here upper approximations from Daniel et al. (2025) as their results are not BFs in general, as it was presented there. In the case of Table 4 (there) $f(a)$ is even more general than pseudo-belief function as the sum of corresponding belief masses over all subsets of Ω is greater than 1. We also skip zero objective (ZO) and Sparsity (SP) geometric approximation, as ZO with its zero objective function returns an ad-hoc BF from the corresponding polytope and SP assigns all belief masses to singletons, thus a large information is added there and the belief structure of pseudo-belief bel_J is completely lost there.

Finally, we have to recall that both correction Algorithms 2 and 3 produce the same results on our simple data example, having negative pseudo-belief masses only on focal elements of cardinality 3. Having our still limited experience with pseudo-belief functions based on Jeffreys confidence interval, we have a hypothesis, the Jeffreys PBFs on $|\Omega| = 4$ have either no negative pseudo-belief masses or have negative belief masses only on focal elements of cardinality 3, hence Algorithms 2 and 3 both produce the same results on any data on any 4-element frame.

Table 2: Comparion pseudo-belief masses derived from Jeffreys intervals $m_J = m_g$ with degree of confidence $\alpha = 0.05$ on cybersecurity data on $|\Omega| = 4$ with its corrections: m_f and m_{f^*} from Daniel et al. (2025),m_{CW} and m_{HD} obtained by geometric Cardinality-Weighted and Dubois-Prade entropy, m_1 and m_2 by correction Algorithms 1 and 2.

A	$bel_J(A)$	$\sum_{B\subsetneq A} m_J(B)$	$m_J(A)$	$m_f(A)$	$m_{f^*}(A)$	$m_{CW}(A)$	$m_{HD}(A)$	$m_1(A)$	$m_2(A)$
$\{\omega_1\}$	0.1635	0	0.16346	0.1635	0.1635	0.1635	0.1739	0.1314	0.1273
$\{\omega_2\}$	0.2114	0	0.21145	0.2114	0.2114	0.2114	0.2219	0.1700	0.1647
$\{\omega_3\}$	0.1792	0	0.17920	0.1792	0.1792	0.1792	0.1897	0.1441	0.1396
$\{\omega_4\}$	0.0496	0	0.04962	0.0496	0.0496	0.0496	0.0603	0.0399	0.0387
$\{\omega_1,\omega_2\}$	0.4606	0.3749	0.08567	0.0682	0.0857	0.0810	0.0647	0.0689	0.0667
$\{\omega_1,\omega_3\}$	0.4226	0.3427	0.07998	0.0648	0.0643	0.0754	0.0590	0.0643	0.0623
$\{\omega_1,\omega_4\}$	0.2615	0.2131	0.04845	0.0367	0.0327	0.0438	0.0273	0.0390	0.0377
$\{\omega_2,\omega_3\}$	0.4798	0.3901	0.08919	0.0756	0.0735	0.0892	0.0683	0.0717	0.0695
$\{\omega_2,\omega_4\}$	0.3135	0.2611	0.05240	0.0422	0.0366	0.0524	0.0313	0.0421	0.0408
$\{\omega_3,\omega_4\}$	0.2786	0.2288	0.04982	0.0420	0.0497	0.0498	0.0287	0.0409	0.0388
$\{\omega_1,\omega_2,\omega_3\}$	0.7776	0.8089	-0.03139	0.0149	0	0	0	0	0.0131
$\{\omega_1,\omega_2,\omega_4\}$	0.5795	0.6111	-0.03158	0.0078	0	0	0	0	0.0022
$\{\omega_1,\omega_3,\omega_4\}$	0.5389	0.5705	-0.03159	0.0031	0	0	0	0	0
$\{\omega_2,\omega_3,\omega_4\}$	0.6001	0.6317	-0.03157	0	0	0	0	0	0.0034
Ω	1.0000	0.8830	0.11690	0.0409	0.0538	0	0.0749	0.1884	0.1952

What can we see in the tables?

The important is that both f and f^* and also both m_1 and m_2 not increase or even decrease their value comparing with $g = bel_J$, i.e., all four are $\leq g$; this corresponds to the fact that f and f^* are lower approximation and bel_i's are constructed using generalized

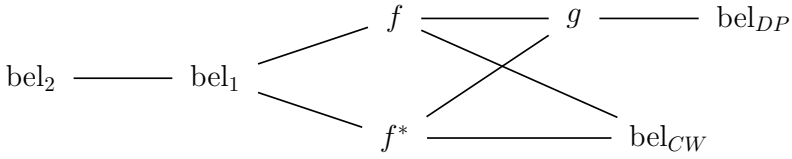
Table 3: Comparison of belief functions — corrected pseudo-beliefs derived from Jeffreys intervals with degree of confidence $\alpha = 0.05$ on cybersecurity data

A	$bel_J(A)$	$\sum_{b \subsetneq a} m_J(A)$	$m_J(A)$	$f(A)$	$f^*(A)$	$bel_{CW}(A)$	$bel_{HD}(A)$	$bel_1(A)$	$bel_2(A)$
$\{\omega_1\}$	0.1635	0.0000	0.16346	0.1635	0.1635	0.1635	0.1739	0.1314	0.1273
$\{\omega_2\}$	0.2114	0.0000	0.21145	0.2114	0.2114	0.2114	0.2219	0.1700	0.1647
$\{\omega_3\}$	0.1792	0.0000	0.17920	0.1792	0.1792	0.1792	0.1897	0.1441	0.1396
$\{\omega_4\}$	0.0496	0.0000	0.04962	0.0496	0.0496	0.0496	0.0603	0.0399	0.0387
$\{\omega_1, \omega_2\}$	0.4606	0.3749	0.08567	0.4431	0.4606	0.4560	0.4606	0.3704	0.3587
$\{\omega_1, \omega_3\}$	0.4226	0.3427	0.07998	0.4075	0.4069	0.4180	0.4226	0.3399	0.3292
$\{\omega_1, \omega_4\}$	0.2615	0.2131	0.04845	0.2498	0.2457	0.2569	0.2615	0.2103	0.2037
$\{\omega_2, \omega_3\}$	0.4798	0.3901	0.08919	0.4663	0.4642	0.4798	0.4798	0.3859	0.3738
$\{\omega_2, \omega_4\}$	0.3135	0.2611	0.05240	0.3033	0.2977	0.3135	0.3135	0.2521	0.2442
$\{\omega_3, \omega_4\}$	0.2786	0.2288	0.04982	0.2708	0.2785	0.2786	0.2786	0.2241	0.2170
$\{\omega_1, \omega_2, \omega_3\}$	0.7776	0.8089	-0.03139	0.7776	0.7776	0.7997	0.7776	0.6505	0.6432
$\{\omega_1, \omega_2, \omega_4\}$	0.5795	0.6111	-0.03158	0.5795	0.5795	0.6018	0.5795	0.4914	0.4782
$\{\omega_1, \omega_3, \omega_4\}$	0.5389	0.5705	-0.03159	0.5389	0.5389	0.5613	0.5389	0.4588	0.4444
$\{\omega_2, \omega_3, \omega_4\}$	0.6001	0.6317	-0.03157	0.6001	0.6001	0.6317	0.6001	0.5080	0.4954
Ω	1.0000	0.8831	0.11690	1.0000	1.0000	0.9954	1.0000	1.0000	1.0000

discounting. Thus all these corrections decreases information of the original PBF g. Moreover, it holds $bel_2 \leq bel_1 \leq f \leq g$ and also $bel_2 \leq bel_1 \leq f^* \leq g$, while f and f^* are mutually $\leq$-incomparable. $bel_2 \leq bel_1$, thus bel_1 is closer to g, nevertheless it has higher computational complexity, which does not play any role on our small example on $|\Omega| = 4$. Both f and f^* are even closer to original g, nevertheless they use ad-hoc negative belief mass redistribution, which may despite closeness to g to add a piece of ad-hoc information, while both bel_1 and bel_2 satisfy all belief proportions at any cardinality of $A \subset \Omega$, hence they better keeps the belief structure of the original PBF $g = bel_J$.

bel_{CW} is $\leq$-incomparable both with bel_{HD} and g while $bel_{HD} \geq g$, thus also $\geq$ all other which are $\leq g$. $bel_{CW} \geq bel_1$, bel_2, f and f^*. Thus both these geometric corrections add some extra information. bel_{CW} has a strange ad-hoc feature: that belief of some couples are increased ($\{\omega_1, \omega_2\}$,$\{\omega_1, \omega_3\}$,$\{\omega_1, \omega_4\}$), while the other keep the same belief mass as the original PBF g has ($\{\omega_2, \omega_3\}$,$\{\omega_2, \omega_4\}$,$\{\omega_3, \omega_4\}$). Hence, there is a challenging open problem: finding more convenient optimization criteria for pseudo-belief correction.

We can summarize our $\leq$-comparison by the following schema.



Unfortunately, negative pseudo-belief masses are relatively quite small and their values

are almost the same (differ on 5th decimal place) in the compared example, thus also the differences of their corrections are rather similar. Hence we have to compare our approaches not only on greater frame of discernment but also on various cases on Ω_4.

Finally, we have to note that this comparison is related only to the one simple case or real data on very small frame of discernment. It is rather a presentation how we can compare our approaches in near future having processed more examples.

One of the interesting open questions is which relations from the $\leq$-comparison schema are general, which of them are frequent, and which of them are rare or even exceptional.

7 Conclusion

Following our preceding contribution Daniel et al. (2025), we have proposed and presented several methods for transforming pseudo-belief functions into classical belief functions. The investigated procedures are based on fundamentally different approaches to correcting pseudo-beliefs. All the presented methods have been compared with those from Daniel et al. (2025) using a simple example based on real cybersecurity data. The implementation of our algorithms is currently under development. This will allow us to perform more comprehensive comparisons on larger frames of discernment and to address several open questions that have emerged in this interesting area of research.

References

R. Bagnara, P. M. Hill, and E. Zaffanella. The parma polyhedra library: Toward a complete set of numerical abstractions for the analysis and verification of hardware and software systems. *Science of Computer Programming*, 72(1-2):3–21, 2008.

F. Cuzzolin. The geometry of consonant belief functions: simplicial complexes of necessity measures. *Fuzzy Sets and Systems*, 161(10):1459–1479, 2010.

F. Cuzzolin. *The geometry of uncertainty: the geometry of imprecise probabilities*. Springer Nature, 2020.

M. Daniel, R. Jiroušek, and V. Kratochvíl. How Sir Harold Jeffreys Would Create a Belief Function Based on Data. In *Proceedings of the 13th Workshop on Uncertainty Processing (WUPES'25)*, pages 92–103, 2025.

D. Dubois and H. Prade. Properties of measures of information in evidence and possibility theories. *Fuzzy sets and systems*, 24(2):161–182, 1987.

G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976. ISBN 9780691100425.

G. M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. ISBN 9780387943657. doi: 10.1007/978-1-4613-8431-1.