Proceedings of the 23rd International Conference on Principles of Knowledge Representation and Reasoning
Main Track

Belief Function Propagation in Quantitative Bipolar Argumentation Frameworks

Jordan Thieyre[1], **Aurélie Beynier**[1], **Sébastien Destercke**[2], **Nicolas Maudet**[1], **Srdjan Vesic**[3]
[1]LIP6, CNRS - Sorbonne Université, Paris, France
[2]Université de technologie de Compiègne, CNRS, Heudiasyc, Compiègne, France
[3]CRIL, CNRS - Univ. Artois, Lens, France
{jordan.thieyre, aurelie.beynier, nicolas.maudet}@lip6.fr, sebastien.destercke@hds.utc.fr, vesic@cril.fr

Abstract

Argumentation theory provides a formal framework to represent and analyse debates where participants propose arguments that attack or support others and assign scores expressing their opinions. Quantitative bipolar argumentation frameworks (*QBAFs*) model such debates by assigning initial weights to arguments and using semantics to compute final acceptability degree that reflect attackers' and supporters' influence. One of the major challenges is setting appropriate initial weights when debaters' opinions are uncertain.

In this paper, we introduce a formal approach to uncertainty propagation in *QBAFs* by representing initial weights as mass functions over a discretized unit interval. We introduce two new propagation models: (i) an exact model that computes final mass functions by combining focal elements of initial weights with parent arguments via bipolar gradual semantics; (ii) a practical approximation that projects the exact mass onto a user-specified partition and reconstructs masses using the Möbius inverse. We prove mathematical properties of the projection and show that the baseline, while computationally efficient, can be overconfident by failing to preserve expectations. Our approximation reduces the exponential complexity of the exact model while satisfying *Epistemic Cautiousness*, yielding acceptability intervals that contain true theoretical expectations and balancing tractability with theoretical soundness.

1 Introduction

Online deliberation platforms allow large number of users to exchange arguments on a given topic, as well as vote (or like) to express agreement or disagreement over specific statements (Shortall et al., 2022; Goldberg et al., 2024). Such systems are typically modelled as argumentation systems, whose exact features may vary depending on the expressivity allowed on the considered platform (Rago and Toni, 2017; Bikakis et al., 2025). One prominent example is given by quantitative bipolar argumentation frameworks (*QBAFs*) (Potyka, 2019; Baroni, Rago, and Toni, 2019), which match quite precisely the Kialo platform (https://www.kialo.com/), for instance.

These platforms raise a number of opportunities and challenges (Shortall et al., 2022). Combining votes and arguments allows for instance to select coherent sets of arguments (*extensions*) which are representative of the collective, a key feature for massive deliberation (Bernreiter et al., 2024; Rossie et al., 2024). Taking a more *gradual* perspective (Amgoud and Ben-Naim, 2018b), the central question of providing principled methods for computing acceptability degrees of arguments has motivated many different proposals of so-called gradual semantics (Amgoud and Ben-Naim, 2018a). While the motivation of online deliberation platforms provide a compelling justification to assign different initial weights to arguments –as votes may reflect different degrees of approval, see *e.g.* Leite and Martins (2011)–, a well-known challenge in this context is that voting is typically very sparse (Goldberg et al., 2024). This means that these methods should account for the inherent uncertainty of the available input. This raises in turn a computational challenge since the propagation of initial uncertain weights can turn out to be expensive. A recent proposal (Thieyre et al., 2025) makes some first steps in addressing this issue: an interval-based representation of uncertainty is put forward, together with a computationally efficient propagation method. However, as we shall see later, it is a rather crude representation of uncertainty which may violate some desirable property.

In this paper we propose an alternative method for handling uncertainty in *QBAFs* based on *belief functions* (Section 3), which provides a more expressive and principled-approach to handle uncertainty. We support this claim by showing that the method of Thieyre et al. (2025) does not satisfy the property of *Epistemic Cautiousness* in the sense that it offers no guarantee regarding its final acceptability degrees when compared to the true values (Section 4). While computing exactly the final acceptability degree is computationally very demanding, we propose an approximate method which limits the computational complexity, while preserving the theoretical properties. We finally run (Section 5) an extensive experimental campaign to evaluate (i) whether the theoretical computational gains are indeed observed in practice, and (ii) how the methods compare in terms of precision of the results, beyond their theoretical guarantees. Interestingly, we show that the conclusions vary depending on the semantics considered, drawing a nuanced picture which justifies the need for a careful assessment of these methods. We focus on acyclic graphs for both theoretical and practical reasons. From a theoretical perspective, most existing bipolar gradual semantics are formulated for acyclic graphs and do not guarantee convergence in the presence of cycles. From a practical standpoint, our goal is to model debates on online platforms and, to the best of our knowledge, such platforms are typically designed so that their underlying argument graphs remain acyclic.

Full code and data are available on our Git[1] repository.

2 Background

Dung's abstract argumentation framework (Dung, 1995) provides a general formalism for reasoning with interacting arguments. In this setting, arguments are represented as nodes of a directed graph, and directed edges encode attack relations. The framework deliberately abstracts away from the internal structure of arguments and focuses solely on their interaction patterns. A substantial body of work has extended this model by explicitly representing supportive relations alongside attacks, giving rise to bipolar argumentation frameworks (Cohen et al., 2014, 2015; Yu et al., 2023).

2.1 Quantitative Bipolar Argumentation Frameworks

Quantitative bipolar argumentation frameworks (*QBAFs*) refine bipolar argumentation by attaching numerical values to arguments. In addition to attack and support relations, each argument is associated with an initial weight in the unit interval $[0, 1]$. This value may capture prior confidence in the argument, an initial bias towards its acceptability, or an aggregation of external evaluations such as user votes, independently of the argument's interactions with others.

Definition 1. *A Quantitative Bipolar Argumentation Framework (QBAF) is a tuple $(\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ where: (i) $\mathcal{A}$ is a finite set of arguments; (ii) $\mathcal{R}^- \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic attack relation; (iii) $\mathcal{R}^+ \subseteq \mathcal{A} \times \mathcal{A}$ is an acyclic support relation; (iv) $w : \mathcal{A} \to [0, 1]$ assigns to each argument a an initial weight w_a.*

For an argument $a \in \mathcal{A}$, the sets $\mathcal{R}_a^-$ and $\mathcal{R}_a^+$ denote its direct attackers and direct supporters, respectively.

2.2 Bipolar Gradual Semantics

Bipolar gradual semantics specify how to compute a final acceptability degree for each argument in a *QBAF*. The resulting value, denoted d_a, is obtained by combining the argument's initial weight with the final acceptability degrees of its attackers and supporters.

Definition 2 (Bipolar Gradual Semantics)**.** *A bipolar gradual semantics s is a function such that, for every argument a,*

$$d_a = s\left(w_a, (d_b)_{b \in \mathcal{R}_a^-}, (d_c)_{c \in \mathcal{R}_a^+}\right).$$

where $(d_b)_{b \in \mathcal{R}_a^-}$ (resp. $(d_c)_{c \in \mathcal{R}_a^+}$) denotes the tuple of final acceptability degrees of attackers (resp. supporters) of a.

In this work, we focus on four well-established bipolar gradual semantics: Quantitative Argumentation Debate (*QuAD*) (Baroni et al., 2015), Discontinuity-Free QuAD (*DF-QuAD*) (Rago et al., 2016), the Exponent-based semantics (*Exb*) (Amgoud and Ben-Naim, 2018a), and the Quadratic Energy Model (*QuEM*) (Potyka, 2018). Since the argumentation graphs considered here are acyclic, final acceptability degrees can be computed by evaluating arguments in a topological order.

The *QuAD* semantics was introduced to assess the relative strength of competing alternatives in Issue-Based Information System (IBIS) style debates. It computes the final acceptability degree of an argument by aggregating its initial weight with the influence exerted by its attackers and supporters.

Definition 3 (*QuAD*)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a QBAF. Let nil represent an ineffective sequence i.e. it is empty or it contains only zeros. The final acceptability degree of a is defined as:*

$$d_a = \begin{cases} f_a^- & \textit{if } (d_b)_{b \in \mathcal{R}_a^-} \neq \textit{nil} \wedge (d_c)_{c \in \mathcal{R}_a^+} = \textit{nil}, \\ f_a^+ & \textit{if } (d_b)_{b \in \mathcal{R}_a^-} = \textit{nil} \wedge (d_c)_{c \in \mathcal{R}_a^+} \neq \textit{nil}, \\ w_a & \textit{if } (d_b)_{b \in \mathcal{R}_a^-} = \textit{nil} \wedge (d_c)_{c \in \mathcal{R}_a^+} = \textit{nil}, \\ \frac{f_a^- + f_a^+}{2} & \textit{otherwise}, \end{cases}$$

where

$$f_a^- = w_a \prod_{b \in \mathcal{R}_a^-} (1 - d_b), \quad f_a^+ = 1 - (1 - w_a) \prod_{b \in \mathcal{R}_a^+} (1 - d_b).$$

The *DF-QuAD* was proposed to address the sensitivity of *QuAD* to weak attacks, which may otherwise cause disproportionate decreases in acceptability.

Definition 4 (*DF-QuAD*)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a QBAF. The final acceptability degree of a is given by:*

$$d_a = \begin{cases} w_a - w_a |f_a^+ - f_a^-| & \textit{if } f_a^- \geqslant f_a^+, \\ w_a + (1 - w_a) |f_a^+ - f_a^-| & \textit{otherwise}, \end{cases}$$

where

$$f_a^- = 1 - \prod_{b \in \mathcal{R}_a^-} (1 - d_b), \quad f_a^+ = 1 - \prod_{b \in \mathcal{R}_a^+} (1 - d_b).$$

The Exponent-based semantics was introduced to overcome the discontinuity of *QuAD* and to mitigate both saturation effects and large amplification effects observed in earlier semantics.

Definition 5 (*Exb*)**.** *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a QBAF. The final acceptability degree of a is:*

$$d_a = 1 - \frac{1 - {w_a}^2}{1 + w_a 2^{E_a}},$$

where

$$E_a = \sum_{c \in \mathcal{R}_a^+} d_c - \sum_{b \in \mathcal{R}_a^-} d_b.$$

While *Exb* avoids the artifacts identified for *QuAD* and *DF-QuAD*, it does not treat attack and support relations symmetrically.

The Quadratic Energy Model (*QuEM*) formulates argument evaluation as a continuous dynamical system in which argument strengths are updated in parallel until convergence. For acyclic frameworks, the resulting equilibrium is unique and can be computed efficiently via a discrete iteration scheme, in linear time.

[1]https://gitlab.com/jthieyre/belief-function-propagtion-in-qbafs

Definition 6 (*QuEM*)**.** *Let $h : \mathbb{R} \to [0,1]$ be defined as*

$$h(e) = \frac{\max(e,0)^2}{1+\max(e,0)^2}.$$

The final acceptability degree of an argument a is:

$$d_a = \begin{cases} w_a + (1-w_a)\, h(E_a) & \text{if } E_a > 0, \\ w_a - w_a\, h(-E_a) & \text{otherwise}, \end{cases}$$

where E_a is the same as for Exb.

2.3 Belief Functions

Representing vote uncertainty requires both to model variability (in the votes), and various quantities of knowledge as the number of votes differ across arguments[2]. It has been argued (Walley, 1991, Ch.5.) that precise probabilities alone cannot really capture these two aspects. One of those arguments is that the representation of ignorance, the uniform distribution, is not invariant under non-linear transformation of a variable, and that a representation of ignorance should have such a property. Other can be found in (Walley, 1991, Ch.5.). Within possible representations solving this issue, belief functions (Shafer, 1976) are one of the simplest models that encompass probabilities and sets under a common umbrella.

Given a space Ω of possible values (in our case, the unit interval $[0,1]$ or a partition of it), a belief measure $Bel : 2^\Omega \to [0,1]$ and its dual plausibility measure Pl are derived from a mass function $m : 2^\Omega \to [0,1]$ using the following formulas:

$$Bel(A) = \sum_{X \subseteq \Omega} m(X) \inf_{x \in X} \mathbb{I}(\{x \in A\}), \tag{1}$$

$$Pl(A) = 1 - Bel(A^c) = \sum_{X \subseteq \Omega} m(X) \sup_{x \in X} \mathbb{I}(\{x \in A\}), \tag{2}$$

with $\mathbb{I}(E)$ the indicator function of an event E, and A^c the complement of A. $Bel(A)$ therefore measures how many elements of evidence implies A, and $Pl(A)$ how many are consistent with A. Sets are captured by the model $m(X) = 1$, and probabilities by having $m(X) = 0$ for all X with $|X| \neq 1$. Note that we will only consider mass functions that are positive on a finite number of intervals or elements.

Belief and plausibility functions can define a set of probabilities $\mathcal{P}_m = \{P : P(A) \geqslant Bel(A) \forall A \subseteq \Omega\}$, in which case a common inferential tool is to consider functions $f : \Omega \to \mathbb{R}$ and to compute lower/upper expectations of f over $\mathcal{P}_m$ defined as:

$$\underline{\mathbb{E}}(f) = \sum_{X \subseteq \Omega} m(X) \inf_{x \in X} f(x) \; ; \; \overline{\mathbb{E}}(f) = \sum_{X \subseteq \Omega} m(X) \sup_{x \in X} f(x).$$

One can see that equations (1) and (2) are instances of those equations when f is the indicator of an event A. Assuming that X are sub-intervals of the unit intervals (as will be the case later on), choosing $f(x) = x$ corresponds to lower/upper expectation of the unknown numerical quantity. In the same vein, using $\mathbb{I}(\{x \leqslant \tau\})$ gives us the upper and lower quantiles for the value τ.

[2]Those two aspects are often distinguished as aleatoric versus epistemic uncertainty.

3 Models

We consider *QBAFs* in which each argument $a \in \mathcal{A}$ is associated with an uncertain initial weight w_a taking values in the unit interval $[0,1]$. To represent this uncertainty, we use mass functions defined over a finite discretisation of $[0,1]$. Let $\Omega = [0, \frac{1}{L}[, [\frac{1}{L}, \frac{2}{L}[, \ldots, [\frac{L-1}{L}, 1]$ denote a partition of the unit interval into L subintervals of equal length. For each argument a, the uncertainty on its initial weight w_a is modelled by a mass function $m_a^i : 2^\Omega \to [0,1]$, satisfying $\sum_{X \subseteq \Omega} m_a^i(X) = 1$. Intuitively, this mass function encodes our knowledge about the possible values of w_a: for any subset $X \subseteq \Omega$, the quantity $m_a^i(X)$ represents the mass assigned to the event $w_a \in X$. Each such subset X is called a *focal element* and correspond to sets of intervals in which the initial weight may lie, allowing the model to capture both imprecision (through intervals) and partial knowledge (through mass assignments to sets of intervals). Given a bipolar gradual semantics s, which combines the initial weight of an argument with the final acceptability degrees of its attackers and supporters, our objective is to propagate these initial mass functions through the argumentation graph so as to obtain, for each argument a, a mass function over its final acceptability degree. The following subsections introduce three propagation models that differ in how uncertainty is handled during this process.

3.1 Baseline Model

The baseline model follows the interval-based propagation scheme introduced in Thieyre et al. (2025). Its purpose is to provide a simple and computationally efficient reference for handling uncertainty in *QBAFs*. The key idea is to summarize the uncertainty associated with each argument by an interval of expected values, and to propagate these expectations through the argumentation graph using a bipolar gradual semantics.

For each argument $a \in \mathcal{A}$, uncertainty on the initial weight is captured by the interval of expectations $\mathbb{E}(w_a) = [\underline{\mathbb{E}}(w_a), \overline{\mathbb{E}}(w_a)]$, defined as

$$\underline{\mathbb{E}}(w_a) = \sum_{X \subseteq \Omega} m_a^i(X) \inf_{x \in X} x \; ; \; \overline{\mathbb{E}}(w_a) = \sum_{X \subseteq \Omega} m_a^i(X) \sup_{x \in X} x$$

These bounds provide a coarse summary of the belief held over the initial weight of a.

The propagation of uncertainty is then performed directly at the level of these expectations. For each argument a, the final acceptability degree in the baseline model is represented by an interval $D_a^B = [\underline{d_a^B}, \overline{d_a^B}]$, obtained by applying the bipolar gradual semantics s to extremal configurations of attackers and supporters:

$$\underline{d_a^B} = s\left(\underline{\mathbb{E}}(w_a), (\overline{d_b^B})_{b \in \mathcal{R}_a^-}, (\underline{d_c^B})_{c \in \mathcal{R}_a^+}\right),$$
$$\overline{d_a^B} = s\left(\overline{\mathbb{E}}(w_a), (\underline{d_b^B})_{b \in \mathcal{R}_a^-}, (\overline{d_c^B})_{c \in \mathcal{R}_a^+}\right).$$

The lower (resp. upper) bound of D_a^B is thus obtained under a pessimistic (resp. optimistic) configuration, where a's initial weight takes its minimal (resp. maximal) expected value, attackers take their maximal (resp. minimal) values, and supporters take their minimal (resp. maximal) values.

Example 1. *Consider a QBAF composed of two arguments a and b, where b attacks a. The initial weights of the arguments are represented by mass functions defined over 2^Ω with $L = 5$ as follows (due to space restriction, only focal elements of the mass functions of the initial weights with non-zero masses are reported i.e. not all elements of 2^Ω):*

X	$[.2, .4[$	$[.4, .6[$	$[.6, .8[$	$[.8, 1]$	$[0, 1]$
$m_a^i(X)$		.14	.29	.29	.28
$m_b^i(X)$	.125	.375	.25		.25

Table 1: Initial mass functions of arguments a and b

The expected initial weight of a is given by:

$$\underline{\mathbb{E}}(w_a) = .4 \times .14 + .6 \times .29 + .8 \times .29 + 0 \times .28 \approx .46,$$
$$\overline{\mathbb{E}}(w_a) = .6 \times .14 + .8 \times .29 + 1 \times .29 + 1 \times .28 \approx .89.$$

Similarly, $\mathbb{E}(w_b) = [.325, .725]$.

Using the QuEM semantics, the baseline interval for the final acceptability degree of: (i) b is the same as the initial weight, since b has no attacker or supporter; (ii) a is computed by propagating pessimistic and optimistic configurations. Since a has no supporters, we obtain:

$$\underline{d_a^B} = s\,(.46, (.725), \varnothing) \approx .30 \quad ; \overline{d_a^B} = s\,(.89, (.325), \varnothing) \approx .80$$

Let $\mathcal{C}(s)$ denote the algorithmic complexity of computing the final acceptability degrees of an argument a using semantics s. Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a *QBAF*. From Thieyre et al. (2025), we know that computing the final acceptability degree of all the arguments in $\mathcal{F}$ using the baseline model is in the complexity of $\mathcal{O}\left(|\mathcal{A}|\mathcal{C}(s)\right)$.

This baseline model is appealing due to its low computational cost and ease of implementation, and it serves as a natural point of comparison for more refined uncertainty propagation methods. However, because it propagates expectations rather than full mass functions, it does not preserve expectations under non-linear gradual semantics, i.e. when the output of s is not a linear combination of its inputs, the image of the expected values generally differs from the expected value of the image. As a consequence, aggregating uncertainty before propagation is not equivalent to propagating uncertainty before aggregation, which may lead the resulting interval D_a^B to under-estimate, over-estimate, or only partially overlap with the expected final acceptability degree produced by the exact model, depending on the chosen semantics, the configuration of attackers and supporters, and the initial weights. Moreover, by reducing uncertainty to intervals of expected values, the baseline model discards much of the information contained in the underlying belief distribution, thereby preventing more expressive forms of epistemic analysis, such as reasoning about distributional shape, medians or quantiles, or expectation bounds associated with alternative utility functions.

3.2 Exact Model

We now formalize the exact propagation of mass functions. Let m_a^f denote the mass function associated with the final acceptability degree of argument a. We assume that the initial mass functions of all arguments are mutually independent. This assumption yields a key tractability advantage: focal elements of final acceptability degree mass functions can be generated without joint dependencies, thereby avoiding an exponential blow-up. Under this assumption, the focal elements of m_a^f are obtained by combining the focal elements of a's attackers, supporters, and its own initial weight through a Cartesian product. While this independence hypothesis is reasonable in settings where the sources of votes or evidence are unknown or heterogeneous (e.g. online platforms), it restricts the model by ignoring potential correlations between arguments. Relaxing it is in principle feasible by introducing joint mass functions over sets of arguments, but this would require modeling and propagating multivariate dependencies, significantly increasing computational complexity (Fetz and Oberguggenberger, 2004).

For each $a \in \mathcal{A}$, let Φ_a denote the set of focal elements of the final mass function m_a^f. Each focal element $F \in \Phi_a$ represents a possible interval in which the final acceptability degree of a may lie, obtained by propagating the uncertainty from a's initial weight together with the final acceptability degrees of its attackers and supporters.

The construction of Φ_a proceeds by considering all possible combinations of input focal elements. To characterize all possible intervals for the final acceptability degree of a, we consider every combination consisting of: (i) a focal element $X \subseteq \Omega$ for a's initial weight, (ii) a tuple $\boldsymbol{F}^- \in \times_{b \in \mathcal{R}_a^-} \Phi_b$ of focal elements for the attackers' final acceptability degrees, and (iii) a tuple $\boldsymbol{F}^+ \in \times_{c \in \mathcal{R}_a^+} \Phi_c$ for the supporters. Each such combination defines a Cartesian product of intervals, whose image through the bipolar gradual semantics s yields a candidate interval for the final acceptability degree of a. Since the inputs of s are intervals, this image is itself an interval. Its bounds are obtained by evaluating s on extremal (pessimistic and optimistic) configurations of the inputs. This construction relies on Lemma 1 of Thieyre et al. (2025), which ensures that, for the considered semantics s, the extrema of the image of a Cartesian product of intervals are attained at the extrema of the input intervals. This property follows from the monotonicity of s with respect to the initial weight, attackers, and supporters.

Formally, for each combination of input focal elements, the resulting focal element $F \in \Phi_a$ is defined as:

$$F = \left[s\left(\underline{X}, \overline{\boldsymbol{F}^-}, \underline{\boldsymbol{F}^+}\right), s\left(\overline{X}, \underline{\boldsymbol{F}^-}, \overline{\boldsymbol{F}^+}\right) \right].$$

where, for $X \subseteq \Omega$, $\underline{X} = \inf_{x \in X} x$ and $\overline{X} = \sup_{x \in X} x$. For $\boldsymbol{F}^- \in \times_{b \in \mathcal{R}_a^-} \Phi_b$ and $\boldsymbol{F}^+ \in \times_{c \in \mathcal{R}_a^+} \Phi_c$, the tuples $\underline{\boldsymbol{F}^-}$ and $\underline{\boldsymbol{F}^+}$ (resp. $\overline{\boldsymbol{F}^-}$ and $\overline{\boldsymbol{F}^+}$) collect the lower (resp. upper) bounds of the corresponding focal elements. A detailed application is illustrated in Example 2.

The exact mass assigned to an output interval $F \in \Phi_a$ is obtained by combining the masses of all input focal elements that produce F. Under the independence assumption, the mass of a given combination $(I, \boldsymbol{J}, \boldsymbol{K})$ is the product of the corresponding input masses. Since distinct combinations may yield the same output interval, the total mass assigned to F is obtained by summing over all such combinations:

$$m_a^f(F) = \sum_{\substack{I,\boldsymbol{J},\boldsymbol{K} \\ s(I,\boldsymbol{J},\boldsymbol{K})\to F}} m_a^i(I) \prod_{b\in\mathcal{R}_a^-} m_b^f(\boldsymbol{J}_b) \prod_{c\in\mathcal{R}_a^+} m_c^f(\boldsymbol{K}_c).$$

Here, $\boldsymbol{J}_b$ (resp. $\boldsymbol{K}_c$) denotes the focal element of $\boldsymbol{J}$ (resp. $\boldsymbol{K}$) associated with argument b (resp. c). The summation aggregates the masses of all combinations of input focal elements whose image under s coincides with F (denoted $s(I, \boldsymbol{J}, \boldsymbol{K}) \to F$), which is necessary since s is commutative with respect to attackers and supporters (considered separately).

Finally, from the exact mass function m_a^f, we derive an interval of expected final acceptability degree, $D_a^E = [\underline{\mathbb{E}}(d_a^E), \overline{\mathbb{E}}(d_a^E)]$, such that:

$$\underline{\mathbb{E}}(d_a^E) = \sum_{F\in\Phi_a} m_a^f(F) \inf_{x\in F} x \ ; \ \overline{\mathbb{E}}(d_a^E) = \sum_{F\in\Phi_a} m_a^f(F) \sup_{x\in F} x$$

The interval D_a^E thus represents the exact expectation range induced by the propagated mass function, and serves as a reference against which approximate propagation schemes can be evaluated.

Example 2. *Following Example 1 and Table 1, we compute the* exact *propagation of the mass functions for the same configuration using the QuEM semantics. Since b has no attackers or supporters, its final mass function coincides with its initial one, i.e., $m_b^f = m_b^i$ (see Table 1). Therefore, $\Phi_b = \Omega$.*

The focal elements of m_a^f are obtained by considering all combinations of focal elements from m_a^i and m_b^f. Each combination (X_a, X_b) corresponds to a pair of intervals representing possible values for the initial weight of a and the final acceptability degree of its attacker b. For each such pair, we compute the image of the corresponding Cartesian product through the semantics s (here, QuEM), which yields an interval for the final acceptability degree of a. Because QuEM satisfies the conditions of Lemma 1 in Thieyre et al. (2025), this image is fully characterized by evaluating s on extremal configurations of the inputs.

Concretely, for a pair of input focal elements X_a and X_b, we define:

$$F = \left[s(\underline{X_a}, (\overline{X_b}), \varnothing),\ s(\overline{X_a}, (\underline{X_b}), \varnothing)\right],$$

where $\varnothing$ denotes the empty tuple of supporters of a.

As a running example, consider the combination $X_a = [.4, .6[$ and $X_b = [.2, .4[$. The pessimistic configuration (lowest a, highest attacker b) gives

$$s(\underline{X_a}, (\overline{X_b}), \varnothing) = s(.4, .4, \varnothing) = .34,$$

while the optimistic configuration (highest a, lowest attacker b) gives

$$s(\overline{X_a}, (\underline{X_b}), \varnothing) = s(.6, .2, \varnothing) = .58.$$

Hence, this pair of input focal elements induces the focal element $F = [.34, .58[$.

Under the independence assumption, the mass associated with this focal element is obtained by multiplying the masses of the corresponding input focal elements:

$$m_a^f([.34, .58[) = m_a^i([.4, .6[) \cdot m_b^i([.2, .4[).$$

Using the numerical values from Table 1, we obtain

$$m_a^f([.34, .58[) = .14 \times .125 = .0175.$$

In other words, if the initial weight of a lies in $[.4, .6[$ with mass $.14$ and the final acceptability degree of b lies in $[.2, .4[$ with mass $.125$, then the final acceptability degree of a lies in $[.34, .58[$ with mass $.0175$.

Repeating this procedure for all combinations of focal elements yields the full set Φ_a of possible output intervals, reported in Table 2. The corresponding masses, obtained by multiplying the masses of each input pair, are given in Table 3 (only non-zero masses are reported due to space constraints).

$X_a \backslash X_b$	$[.2, .4[$	$[.4, .6[$	$[.6, .8[$	$[0, 1]$
$[.4, .6[$	$[.34, .58[$	$[.29, .52[$	$[.24, .44[$	$[.2, .6]$
$[.6, .8[$	$[.52, .77[$	$[.44, .69[$	$[.37, .59[$	$[.3, .8]$
$[.8, 1]$	$[.69, .96[$	$[.59, .86[$	$[.49, .73[$	$[.4, 1]$
$[0, 1]$	$[0, .96[$	$[0, .86[$	$[0, .73[$	$[0, 1]$

Table 2: Exact focal elements Φ_a obtained by combining the focal elements of m_a^i and m_b^i using *QuEM* semantics

$X_a \backslash X_b$	$[.2, .4[$	$[.4, .6[$	$[.6, .8[$	$[0, 1]$
$[.4, .6[$	.0175	.0525	.035	.035
$[.6, .8[$	.03625	.10875	.0725	.0725
$[.8, 1]$	.03625	.10875	.0725	.0725
$[0, 1]$	.035	.105	.07	.07

Table 3: Exact final masses of m_a^f obtained by multiplying the masses of m_a^i and m_b^i for each input pair

Proposition 1. *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a QBAF. Computing the final acceptability degree of all the arguments in $\mathcal{F}$ using the exact model is in the complexity of $\mathcal{O}\left(2^{L|\mathcal{A}|+1}\mathcal{C}(s)\right)$.*

Proof. Let $\nu \in \mathbb{N}$ denote the level of an argument in a *QBAF*. An argument a is of level $\nu(a) = 0$ if it is not attacked nor supported. Otherwise, the level is $\nu(a) = 1 + \max\{\nu(b) | b \in \mathcal{R}^-(a) \cup \mathcal{R}^+(a)\}$. Let $n_{anc}(a)$ be the number of ancestors of a. Let $Q(\nu)$ be the following proposition: the number of calls to the semantics s to compute the mass function of the final acceptability degree of an argument a of level $\nu(a)$ is at most $2^{L(n_{anc}(a)+1)+1}$. Let us prove now by induction that, $\forall \nu \in \mathbb{N}$, $Q(\nu)$ is true.

Base : Let $a \in \mathcal{A}$ be an argument of level $\nu(a) = 0$. Following the procedure described above, we call the semantics s twice for each focal element in Φ_a, so we first need to bound $|\Phi_a|$. We know that:

$$|\Phi_a| \leqslant \left| 2^{\Omega} \times \bigtimes_{b \in \mathcal{R}_a^-} \Phi_b \times \bigtimes_{c \in \mathcal{R}_a^+} \Phi_c \right|$$

Since a has neither attacker nor supporter i.e. $n_{anc}(a) = 0$, we have:

$$|\Phi_a| \leqslant \left|2^{\Omega}\right| \iff |\Phi_a| \leqslant 2^L$$

Thus, the number of calls to s is at most $2 \times 2^L = 2^{L+1}$, which, because $n_{anc}(a) = 0$, is equal to $2^{L(n_{anc}(a)+1)+1}$. Hence $Q(0)$ is true.

Step : Let us suppose that $\forall a \in \mathcal{A}$, if $\nu(a) \leqslant \nu_0$, then $Q(\nu(a))$ is true and let us prove that for any $a \in \mathcal{A}$ such that $\nu(a) = \nu_0 + 1$, then $Q(\nu_0 + 1)$ is true. We still have:

$$|\Phi_a| \leqslant \left| 2^{\Omega} \times \bigtimes_{b \in \mathcal{R}_a^-} \Phi_b \times \bigtimes_{c \in \mathcal{R}_a^+} \Phi_c \right|$$

By induction hypothesis, $\forall b \in \mathcal{R}_a^- \cup \mathcal{R}_a^+$, $\nu(b) \leqslant \nu_0$ and so $Q(\nu(b))$ is true. Thus:

$$\begin{aligned} & |\Phi_a| \leqslant 2^L \prod_{b \in \mathcal{R}_a^-} 2^{L(n_{anc}(b)+1)} \prod_{c \in \mathcal{R}_a^+} 2^{L(n_{anc}(c)+1)} \\ \iff & |\Phi_a| \leqslant 2^{L\left(1+\sum_{b\in\mathcal{R}_a^-}(n_{anc}(b)+1)+\sum_{c\in\mathcal{R}_a^+}(n_{anc}(c)+1)\right)} \\ \iff & |\Phi_a| \leqslant 2^{L(n_{anc}(a)+1))} \end{aligned}$$

For each focal element, we need to call s twice, so the total number of calls is at most $2^{L(n_{anc}(a)+1)+1}$. Hence $Q(\nu_0+1)$ is true.

Conclusion : $\forall \nu \in \mathbb{N}$, $Q(\nu)$ is true.

Thus, the total number of calls to the semantics s on the entire graph is:

$$\sum_{a \in \mathcal{A}} 2^{L(n_{anc}(a)+1)+1}$$

This sum is asymptotically dominated by the term corresponding to the central argument, i.e. $2^{L|\mathcal{A}|+1}$. Finally, the complexity of the exact model is then $\mathcal{O}\left(2^{L|\mathcal{A}|+1}\mathcal{C}(s)\right)$. $\square$

3.3 Approximate Model

Since the cardinality of Φ_a grows exponentially with the number of ancestors, we introduce an approximation method that projects the exact mass function onto a fixed target partition $\Pi = \{[0, \frac{1}{n_p}[, [\frac{1}{n_p}, \frac{2}{n_p}[, \ldots, [\frac{n_p-1}{n_p}, 1]\}$ consisting of n_p subintervals of equal length. The exact mass function m_a^f is approximated by a simplified mass function $\widetilde{m}_a^f$ defined over Π.

The approximation proceeds in two steps. First, we compute an approximate belief function $\widetilde{Bel}_a^f$ obtained by projecting the exact mass function onto the partition Π. Then, we recover the corresponding mass function $\widetilde{m}_a^f$ using the Möbius inverse. This approximated mass function is subsequently used as input for the propagation to the children of a. By projecting the mass function onto the fixed partition Π, the approximation constrains the growth of focal elements and prevents the combinatorial explosion observed in the exact model.

This construction is motivated by the duality between belief functions and mass functions. While the projection onto Π naturally yields an outer-approximated belief function $\widetilde{Bel}_a^f$, the propagation mechanism i.e. bipolar gradual semantics, operates on mass functions. The Möbius inversion provides the transformation between these two representations on the discretized space, ensuring that $\widetilde{m}_a^f$ is a valid mass function consistent with the approximated belief function.

First, for any interval $I \subseteq \Pi$, we calculate the belief accumulated in I from the exact mass function:

$$\widetilde{Bel}_a^f(I) = \sum_{\substack{F \in \Phi_a \\ F \subseteq I}} m_a^f(F)$$

Finally, the approximated mass $\widetilde{m}_a^f(I)$ is recovered following the Möbius inverse:

$$\widetilde{m}_a^f(I) = \sum_{B \subseteq I} (-1)^{|I-B|} \widetilde{Bel}_a^f(B)$$

We then extract from the approximated mass function $\widetilde{m}_a^f$ a summary interval of expected final acceptability degrees, in the same spirit as for the exact model. This interval captures the expectation range induced by the approximated mass function defined on the fixed partition Π.

We thus define $D_a^A = [\underline{\mathbb{E}}(d_a^A), \overline{\mathbb{E}}(d_a^A)]$ such that:

$$\underline{\mathbb{E}}(d_a^A) = \sum_{I \subseteq \Pi} \widetilde{m}_a^f(I) \inf_{x \in I} x \ ; \ \overline{\mathbb{E}}(d_a^A) = \sum_{I \subseteq \Pi} \widetilde{m}_a^f(I) \sup_{x \in I} x$$

Example 3. *To reduce computational complexity, we project the exact mass function m_a^f onto a fixed partition $\Pi = \{[0, .2[, [.2, .4[, [.4, .6[, [.6, .8[, [.8, 1]\}$. First, we compute the approximated belief function for each interval $I \subseteq \Pi$. For instance, following Table 2 and 3, if $I = [.2, .6[$, then:*

$$\begin{aligned} \widetilde{Bel}_a^f([.2, .6[) &= m_a^f([.34, .58[) + m_a^f([.29, .52[) \\ &+ m_a^f([.24, .44[) + m_a^f([.2, .6[) \\ &+ m_a^f([.37, .59[) \\ &= .0175 + .0525 + .035 \times 2 + .0725 \\ &= .2125 \end{aligned}$$

Once the beliefs have been calculated (see Table 4 below) for all $I \subseteq \Pi$, the approximate mass function is then recovered using the Möbius inverse:

$$\begin{aligned} \widetilde{m}_a^f([.2, .6[) &= (-1)^1 \widetilde{Bel}_a^f([.2, .4[) + (-1)^1 \widetilde{Bel}_a^f([.4, .6[) \\ &+ (-1)^0 \widetilde{Bel}_a^f([.2, .6[) \\ &= 0 + 0 + .2125 \\ &= .2125 \end{aligned}$$

All the results obtained for $\widetilde{Bel}_a^f$ and $\widetilde{m}_a^f$ are listed in the following Table (due to space restriction, only focal elements *of the approximate mass and belief functions of a's final acceptability degree with non-zero beliefs are reported):*

2^Π	$\widetilde{Bel}_a^f$	$\widetilde{m}_a^f$
$[.2, .6[$	.2125	.2125
$[.4, .8[$	.2175	.2175
$[.6, 1]$	.03625	.03625
$[0, .6[$	.2125	0
$[0, .2[\cup[.6, 1]$	.03625	0
$[.2, .8[$	.5025	.0725
$[.2, .6[\cup[.8, 1]$	.2125	0
$[.2, .4] \cup [.6, 1]$	.03625	0
$[.4, 1]$	.435	.18125
$[0, .8[$	.5725	.07
$[0, .6[\cup[.8, 1]$	.2125	0
$[0, .4] \cup [.6, 1]$	.03625	0
$[0, .2[\cup[.4, 1]$	.435	0
$[.2, 1]$	.72	0
$[0, 1]$	1	.21

Table 4: Approximate beliefs $\widetilde{Bel}_a^f$ and masses $\widetilde{m}_a^f$

Proposition 2. *Let $\mathcal{F} = (\mathcal{A}, \mathcal{R}^-, \mathcal{R}^+, w)$ be a QBAF. Let $\Delta^- = \max_{a\in\mathcal{A}} deg^-(a)$ be the maximum in-degree of $\mathcal{F}$. Computing the final acceptability degree of all the arguments in $\mathcal{F}$ using the approximate model is in the complexity of $\mathcal{O}\left(2^{L+1}{n_p}^{2\Delta^-}\mathcal{C}(s)\right)$.*

Proof. Let $a \in \mathcal{A}$ be an argument. Following the procedure above, we call the semantics s twice for each focal element in Φ_a, so we first need to bound $|\Phi_a|$. We know that:

$$|\Phi_a| \leqslant \left| 2^\Omega \times \bigtimes_{b\in\mathcal{R}_a^-} \widetilde{\Phi}_b \times \bigtimes_{c\in\mathcal{R}_a^+} \widetilde{\Phi}_c \right|$$

where $\widetilde{\Phi}_b$ (resp. $\widetilde{\Phi}_c$) denote the set of focal elements of $\widetilde{m}_b^f$ (resp. $\widetilde{m}_c^f$). However, $\forall b \in \mathcal{R}_a^-$, $\forall c \in \mathcal{R}_a^+$, $|\widetilde{\Phi}_b| = |\widetilde{\Phi}_c| = \frac{n_p(n_p+1)}{2}$ (see Section 4.1). Then,

$$|\Phi_a| \leqslant 2^L \prod_{b\in\mathcal{R}_a^-} \frac{n_p(n_p+1)}{2} \prod_{c\in\mathcal{R}_a^+} \frac{n_p(n_p+1)}{2}$$

$$\iff |\Phi_a| \leqslant 2^L \left(\frac{n_p(n_p+1)}{2}\right)^{|\mathcal{R}_a^-\cup\mathcal{R}_a^+|}$$

Thus, we need to call s at most $2^{L+1}\left(\frac{n_p(n_p+1)}{2}\right)^{|\mathcal{R}_a^-\cup\mathcal{R}_a^+|}$.

Then, computing $\widetilde{Bel}_a^f$ is done by iterating over all elements of Φ_a, and computing $\widetilde{m}_a^f$ by iterating over all elements of Π, so there is no need to make new calls to s. Thus, the total number of calls to the semantics s on the entire graph is:

$$\sum_{a\in\mathcal{A}} 2^{L+1}\left(\frac{n_p(n_p+1)}{2}\right)^{|\mathcal{R}_a^-\cup\mathcal{R}_a^+|}$$

This sum is asymptotically dominated by the term corresponding to the argument with the highest number of attackers and supporters, i.e. $2^{L+1}\left(\frac{n_p(n_p+1)}{2}\right)^{\Delta^-}$. Finally, the complexity of the approximate model is then $\mathcal{O}\left(2^{L+1}{n_p}^{2\Delta^-}\mathcal{C}(s)\right)$. □

Having introduced the three propagation models, we now examine their theoretical properties.

4 Properties

4.1 Properties of the Outer-Approximation

In this section, our goal is to study the mathematical properties of $\widetilde{Bel}_a^f$, the approximation of the initial Bel_a^f on the partitioned space Π and the corresponding σ-algebra. Note that as $\widetilde{Bel}_a^f$ is the restriction of Bel_a^f to 2^Π (i.e., $\widetilde{Bel}_a^f(X) = Bel_a^f(X)$ for all $X \in 2^\Pi$), we are certain that this is a belief function, and also that this is the closest approximation of Bel_a^f on Π. Note also that the partition is to be fixed by the user, and will directly impact the complexity of the obtained representation.

$\widetilde{Bel}_a^f$ has actually a very nice structure leading to a not too complex belief function, akin to the one initially considered by Strat (1984). We first prove that $\widetilde{Bel}_a^f$ is additive over subsets X that are discontinuous.

Proposition 3. *Consider an event $X = \{[\frac{k}{n_p}, \frac{\ell}{n_p}[\cup[\frac{m}{n_p}, \frac{n}{n_p}[\}$ where $k, \ell, m, n \in [\![0, n_p]\!]^4$ and $m \geqslant \ell + 1$, then we have*

$$\widetilde{Bel}_a^f(X) = \widetilde{Bel}_a^f([\frac{k}{n_p}, \frac{\ell}{n_p}[) + \widetilde{Bel}_a^f([\frac{m}{n_p}, \frac{n}{n_p}[)$$

Proof. To state this proof, we need to show that we do not have focal elements of Bel_a^f that is included in $[\frac{k}{n_p}, \frac{\ell}{n_p}[\cup[\frac{m}{n_p}, \frac{n}{n_p}[$ without being included in either $[\frac{k}{n_p}, \frac{\ell}{n_p}[$ or $[\frac{m}{n_p}, \frac{n}{n_p}[$.

In order to prove this, let us consider an initial focal element F, and show that

$$F \subseteq [\frac{k}{n_p}, \frac{\ell}{n_p}[\cup[\frac{m}{n_p}, \frac{n}{n_p}[\implies F \subseteq [\frac{k}{n_p}, \frac{\ell}{n_p}[\oplus F \subseteq [\frac{m}{n_p}, \frac{n}{n_p}[$$

with $\oplus$ symbolising the XOR logical operation. To see this, simply observes that if $F \cap [\frac{k}{n_p}, \frac{\ell}{n_p}[\neq \emptyset$ and $F \cap [\frac{m}{n_p}, \frac{n}{n_p}[\neq \emptyset$, then F should contain all elements between $\frac{\ell}{n_p}$ and $\frac{m}{n_p}$, thus contradicting the fact that $F \subseteq [\frac{k}{n_p}, \frac{\ell}{n_p}[\cup[\frac{m}{n_p}, \frac{n}{n_p}[$. □

This proof can then be extended to any number of subintervals and discontinuities following the same argument, thus proving that $\widetilde{Bel}_a^f$ is additive over sequences of subintervals. This is formalised by the next corollary.

Corollary 1. *Consider an event*

$$X = \{[\frac{n^1}{n_p}, \frac{n^2}{n_p}[\cup \ldots \cup [\frac{n^i}{n_p}, \frac{n^{i+1}}{n_p}[\cup \ldots \cup [\frac{n^{K-1}}{n_p}, \frac{n^K}{n_p}[\}$$

with $n^{2j+1} \geqslant n^{2j}+1$ *for* $j=1,\ldots,K/2$. *Then* $\widetilde{Bel}^f_a(X)=\sum_{j=1}^K \widetilde{Bel}^f_a([\frac{n^j}{n_p},\frac{n^{j+1}}{n_p}[)$.

We can now proceed to prove our next result, that concern not $\widetilde{Bel}^f_a$ but $\widetilde{m}^f_a$, and essentially shows that the obtained mass function also only bears on sets of sequential sub-intervals.

Proposition 4. *The mass* $\widetilde{m}^f_a$ *derived from* $\widetilde{Bel}^f_a$ *through the Möbius inverse is such that* $\widetilde{m}^f_a(X)$ *is positive if and only if* X *is of the form* $[\frac{k}{n_p},\frac{\ell}{n_p}[$*, i.e., has no gap in it.*

Proof. The proof works by contraposition. Assume that we have some positive mass function bearing on a set $X=[\frac{k}{n_p},\frac{\ell}{n_p}[\cup[\frac{m}{n_p},\frac{n}{n_p}[$ as in Proposition 3. Then, obviously $X \nsubseteq [\frac{k}{n_p},\frac{\ell}{n_p}[$ and $X \nsubseteq [\frac{m}{n_p},\frac{n}{n_p}[$, so it is counted neither in the computation $\widetilde{Bel}^f_a([\frac{k}{n_p},\frac{\ell}{n_p}[)$ nor the one $\widetilde{Bel}^f_a([\frac{m}{n_p},\frac{n}{n_p}[)$, but is counted in the computation of $\widetilde{Bel}^f_a([\frac{k}{n_p},\frac{\ell}{n_p}[\cup[\frac{m}{n_p},\frac{n}{n_p}[)$, which means that we would have

$$\widetilde{Bel}^f_a([\frac{k}{n_p},\frac{\ell}{n_p}[\cup[\frac{m}{n_p},\frac{n}{n_p}[)>\widetilde{Bel}^f_a([\frac{k}{n_p},\frac{\ell}{n_p}[)+\widetilde{Bel}^f_a([\frac{m}{n_p},\frac{n}{n_p}[),$$

thus contradicting Proposition 3 that we know to be true. □

Basically, Propositions 3 and 4 both mean that we are guaranteed that, at worst, the number of distinct $\widetilde{Bel}^f_a$ values one needs to compute or the number of focal elements of $\widetilde{m}^f_a$ will be of the order $\frac{n_p(n_p-1)}{2}$, which clearly limits the computational burden when compared to generic belief functions.

4.2 Properties of the Approximate Model

Since the exact model is very difficult to calculate in practice, a desirable property for an approximation model is to avoid overconfidence in the final acceptability degree of an argument, i.e., neither of the two final bounds should lie within the interval of the exact model. This requirement is captured by the notion of Epistemic Cautiousness, which ensures that the approximation does not yield conclusions that are more precise than what is warranted by the available uncertain information. This property is particularly relevant in our setting, where beliefs are constructed from sparse and potentially unreliable inputs, and where overly confident estimates could lead to misleading conclusions. For this reason, we focus on Epistemic Cautiousness as a key criterion to assess the reliability of approximation methods.

Property 1 (Epistemic Cautiousness (*EC*))**.** *An approximation model satisfies EC iff* $D^E_a \subseteq D^*_a$*, where, in our cases,* D^*_a *stands for* D^B_a *or* D^A_a.

Proposition 5. *Approximate model satisfies EC but Baseline model does not.*

Proof for Approximate model. Since $\widetilde{m}^f_a$ is an outer approximation of m^f_a then $\underline{\mathbb{E}}(d^A_a) \leqslant \underline{\mathbb{E}}(d^E_a)$ and $\overline{\mathbb{E}}(d^E_a) \leqslant \overline{\mathbb{E}}(d^A_a)$, and therefore $D^E_a \subseteq D^A_a$. □

Proof for Baseline model. We provide a counter-example using *QuEM* semantics. Consider a *QBAF* composed of three arguments a, b and c, where b attacks a and c supports a. Suppose that $m^i_a([0;.2[)=\frac{1}{2}$, $m^i_a([0;1[)=\frac{1}{2}$, $m^i_b([.8;1[)=\frac{1}{3}$, $m^i_b([0;1[)=\frac{2}{3}$, $m^i_c([.8;1[)=\frac{1}{3}$ and $m^i_c([0;1[)=\frac{2}{3}$. Given that b and c have neither attackers nor supporters, then $m^f_b=m^i_b$ and $m^f_c=m^i_c$.

Following exact propagation detailed in Section 3.2, we have:

F	$[0,\frac{5}{34}[$	$[0,\frac{3}{13}[$	$[0,.6[$	$[0,\frac{25}{34}[$	$[0,1]$
$m^f_a(F)$	$\frac{1}{18}$	$\frac{4}{18}$	$\frac{4}{18}$	$\frac{1}{18}$	$\frac{8}{18}$

Thus, $D^E_a=[0,\frac{13487}{19890}]$.

Following baseline model detailed in Section 3.1, we have $\mathbb{E}(w_a)=[0,\frac{3}{5}]$, $\mathbb{E}(w_b)=[\frac{4}{15},1]$ and $\mathbb{E}(w_c)=[0,\frac{11}{15}]$. Then:

$$\begin{aligned}\underline{d^B_a}&=s\left(\underline{\mathbb{E}}(w_a),(\overline{\mathbb{E}}(w_b)),(\underline{\mathbb{E}}(w_c))\right)=s(0,(1),(0))\\&=0,\\ \overline{d^B_a}&=s\left(\overline{\mathbb{E}}(w_a),(\underline{\mathbb{E}}(w_b)),(\overline{\mathbb{E}}(w_c))\right)=s\left(\frac{3}{5},(\frac{4}{15}),(\frac{11}{15})\right)\\&=\frac{92}{137}.\end{aligned}$$

Thus, $D^B_a=[0,\frac{92}{137}]$, and therefore $D^B_a \subseteq D^E_a$. □

These theoretical results clarify the epistemic guarantees offered by the approximate model and the limitations of the baseline model. However, they do not inform us about the practical impact of these limitations, which motivates an empirical comparison of the baseline, exact and approximate models on real-world argumentation graphs.

5 Evaluation on a Real-World Dataset

We evaluated the three models on argumentation graphs extracted from Kialo, an online platform for structured debates. On Kialo, discussions are organized around a central claim, to which users add supporting or opposing claims ("*PRO*" and "*CON*"). Contributions are curated by moderators to ensure clarity and consistency, resulting in clean argument trees with explicit support and attack relations. Users vote on claims using a discrete scale from 0 to 4. As in Thieyre et al. (2025), we transform these votes into mass functions over initial weights using the Imprecise Dirichlet Model.

The complete dataset contains 2,959 debates, 377,182 claims, and 590,261 votes. Due to the exponential complexity of the exact model, we restricted the analysis to 13,285 arguments for which exact propagation was computationally feasible. Among these, we retained 1,632 non-trivial arguments. The excluded cases correspond to arguments without attackers or supporters and whose initial mass is entirely concentrated on the focal element $[0,1]$. In such situations, the final mass remains on $[0,1]$ for all three models and all four semantics, yielding identical results. Excluding these trivial cases allows us to focus on configurations where the differences between models are meaningful. In all experiments, the approximate model was computed using a partition size $n_p = 5$.

We investigated the following five empirical criteria, reported in Table 5:

- $\cancel{\mathbf{EC}}$: the baseline model does not satisfy *EC*;
- $\mathbf{B}{\geqslant}\mathbf{A}$: the baseline model is a better or equal outer approximation of the exact model than the approximate model;
- $\mathbf{B}{>}\mathbf{A}$: the baseline model is strictly a better outer approximation of the exact model than the approximate model;
- $\mathbf{B}{\trianglerighteq}\mathbf{A}$: the baseline model is closer or equally close (in absolute value) to the exact model than the approximate model;
- $\mathbf{B}{\triangleright}\mathbf{A}$: the baseline model is strictly closer (in absolute value) to the exact model than the approximate model.

Here, "closer in absolute value" means that the lower bound of D_a^B is closer (in absolute difference) to the lower bound of D_a^E than the lower bound of D_a^A, and similarly for the upper bounds. A "better or equal outer approximation" additionally requires that $D_a^E \subseteq D_a^B \subseteq D_a^A$, while a "strictly better outer approximation" requires $D_a^E \subseteq D_a^B \subset D_a^A$.

	QuAD	*DF-QuAD*	*Exb*	*QuEM*
$\cancel{\mathbf{EC}}$	187 11.5%	2 0.1%	206 12.6%	39 2.4%
$\mathbf{B}{\geqslant}\mathbf{A}$	1344 82.3%	1403 86.0%	457 28.0%	1447 88.7%
$\mathbf{B}{>}\mathbf{A}$	20 1.2%	46 2.8%	77 4.7%	185 11.3%
$\mathbf{B}{\trianglerighteq}\mathbf{A}$	1503 92.1%	1616 99.0%	544 33.3%	1559 95.5%
$\mathbf{B}{\triangleright}\mathbf{A}$	124 7.6%	165 10.1%	157 9.6%	224 13.7%

Table 5: Empirical comparison between models

As expected from the theoretical analysis, the baseline model violates *EC* in a non-negligible number of cases, particularly under *QuAD* and *Exb*. However, the frequency of violations varies considerably across semantics, being almost negligible under *DF-QuAD* and relatively limited under *QuEM*. This confirms that epistemic overconfidence is sensitive to the aggregation behavior induced by the underlying semantics.

When considering outer-approximation quality, the results show that under *QuAD*, *DF-QuAD*, and *QuEM*, the baseline model is a better or equal outer approximation than the approximate model in a large majority of cases (between 82.3% and 88.7%). In contrast, under *Exb*, this holds in only 28.0% of the cases. Strict improvements ($\mathbf{B}{>}\mathbf{A}$) remain relatively rare overall, but are more frequent under *QuEM*. These results indicate that, although the approximate model is guaranteed to satisfy *EC*, it does not systematically yield tighter outer approximations than the baseline model in practice.

A similar pattern emerges when considering numerical closeness to the exact model. For *QuAD*, *DF-QuAD*, and *QuEM*, the baseline model is closer or equally close to the exact interval in at least 92.1% of the non-trivial cases. This suggests that, despite its theoretical limitations, the baseline model often provides a numerically accurate approximation of expected final acceptability degrees under these semantics. One possible explanation is that, for many argumentation configurations encountered in practice, these semantics exhibit near-linear behavior, thereby mitigating the discrepancy between summarize-then-propagate and propagate-then-summarize strategies.

This behavior does not extend to *Exb*, for which both the outer-approximation and closeness criteria show markedly weaker performance of the baseline model. The stronger non-linear and asymmetric effects induced by this semantics appear to amplify the information loss caused by early summarization of uncertainty.

Overall, the results reveal a clear trade-off. The approximate model provides principled epistemic guarantees through *EC*, while the baseline model often delivers competitive or even superior approximations in practice under several widely used semantics. The choice between models should therefore depend on whether epistemic soundness or empirical tightness of approximation is the primary objective.

6 Related Work and Conclusion

There exists a significant body of work at the interface of uncertainty and argumentation, in particular probabilistic (Hunter et al., 2021), fuzzy (Janssen, Vermeir, and De Cock, 2008), and epistemic (Hunter and Thimm, 2014) approaches.

Among these approaches, probabilistic argumentation has received particular attention. Early and subsequent works focus on modeling uncertainty either over the structure of the argument graph or over the acceptance of arguments themselves, often assuming precise probability values (Dung, 1995; Hunter et al., 2021). However, in many real-world applications, probabilities originate from heterogeneous sources, expert judgments, or incomplete data, making it difficult to justify precise numerical assignments. This limitation has motivated the development of frameworks based on imprecise probabilities, which allow uncertainty to be represented by sets of probability distributions rather than a single one.

Several works have explored the use of imprecise probabilistic reasoning within abstract argumentation frameworks (Baroni, Giacomin, and Vicig, 2014; Thimm et al., 2018). These approaches relax the assumption of precise probabilities and provide more robust epistemic interpretations, sometimes in combination with probabilistic logic to enhance expressive power (Hunter and Potyka, 2023). Nevertheless, these contributions are typically confined to Dung-style abstract argumentation frameworks and extension-based semantics, and do not address gradual acceptability or bipolar interactions.

It is also possible to combine abstract argumentation for attack relations with credal networks for modeling supports, as proposed by Morveli-Espinoza, Nieves, and Tacla (2023b) in their credal bipolar argumentation framework. In this approach, each argument is associated with a credal set (i.e., a set of probability distributions), and support relations are explicitly interpreted as causal dependencies modeled via credal networks, yielding conditional credal sets that depend on the presence of supporting arguments. Uncertainty is thus propagated through the argument graph by computing lower and upper bounds over these sets, allowing one to capture epistemic imprecision arising from multiple sources without committing to a single probability distribution. However, reasoning remains tied to extension-based semantics, where the uncertainty of extensions is derived from aggregating the credal information of their constituent arguments. The same authors later introduced a gradual semantics grounded in credal network theory (Morveli Espinoza, Nieves, and Tacla, 2023a), specifically designed for support argumentation frameworks.

In contrast to these works, the present paper focuses on uncertainty propagation in quantitative bipolar argumentation frameworks under gradual semantics, using belief functions to represent imprecise information about argument weights. We investigate how uncertainty can be propagated through existing bipolar gradual semantics and analyse the theoretical and empirical properties of different propagation strategies. In particular, we compare three propagation models and formalize the Epistemic Cautiousness property, showing that while the approximate model satisfies this property by construction, the baseline model does not in general. More broadly, our work highlights the trade-offs between epistemic guarantees and practical approximation quality.

Beyond this theoretical distinction, our empirical evaluation on large-scale real-world argumentation graphs extracted from Kialo revealed a more nuanced picture. Although the approximate model provides stronger epistemic guarantees, the baseline model often yields numerically tighter approximations of the exact model under several semantics, notably *QuAD*, *DF-QuAD*, and *QuEM*. These results highlight a fundamental trade-off between epistemic soundness and practical approximation quality, suggesting that the choice of propagation model should depend on both the semantics in use and the intended application.

Several directions for future work naturally follow from this study. In particular, an important question is whether the approximation step itself behaves in a manner that is coherent with the underlying mechanisms of argumentation. Investigating this issue could help determine to what extent approximate propagation preserves the intended influence of supports and attacks, and could guide the design of approximation methods that better balance epistemic guarantees with argumentative fidelity.

Acknowledgments

This work has been supported by the ANR project AGGREEY (ANR-22-CE23-0005), and under the France 2030 program by the project AI Cluster PostGenAI (ANR-23-IACL-0007).

AI Declaration

The authors employed Google Gemini 3 Pro to optimize the execution speed of the Möbius inverse and to reduce the number of semantics calls by leveraging the commutativity property of the underlying semantics.

References

Amgoud, L., and Ben-Naim, J. 2018a. Evaluation of arguments in weighted bipolar graphs. *International Journal of Approximate Reasoning* 99:39–55.

Amgoud, L., and Ben-Naim, J. 2018b. Weighted bipolar argumentation graphs: Axioms and semantics. In *Twenty-Seventh International Joint Conference on Artificial Intelligence-IJCAI 2018*, 5194–5198.

Baroni, P.; Romano, M.; Toni, F.; Aurisicchio, M.; and Bertanza, G. 2015. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation* 6(1):24–49.

Baroni, P.; Giacomin, M.; and Vicig, P. 2014. On Rationality Conditions for Epistemic Probabilities in Abstract Argumentation. In *Computational Models of Argument*. IOS Press. 121–132.

Baroni, P.; Rago, A.; and Toni, F. 2019. From fine-grained properties to broad principles for gradual argumentation: A principled spectrum. *International Journal of Approximate Reasoning* 105:252–286.

Bernreiter, M.; Maly, J.; Nardi, O.; and Woltran, S. 2024. Combining voting and abstract argumentation to understand online discussions. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '24, 170–179. Richland, SC: IFAAMAS.

Bikakis, A.; Flouris, G.; Leite, J.; and Patkos, T. 2025. Social argumentation systems. *Journal of Applied Logics: IfCoLog Journal of Logics and their Applications* 12(3):251–294.

Cohen, A.; Gottifredi, S.; García, A. J.; and Simari, G. R. 2014. A survey of different approaches to support in argumentation systems. *The Knowledge Engineering Review* 29(5):513–550.

Cohen, A.; Gottifredi, S.; García, A. J.; and Simari, G. R. 2015. An approach to abstract argumentation with recursive attack and support. *Journal of Applied Logic* 13(4):509–533.

Dung, P. M. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence* 77(2):321–357.

Fetz, Th., and Oberguggenberger, M. 2004. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering & System Safety* 85(1):73–87.

Goldberg, B.; Acosta-Navas, D.; Bakker, M.; Beacock, I.; Botvinick, M.; Buch, P.; DiResta, R.; Donthi, N.; Fast, N.; Iyer, R.; Jalan, Z.; Konya, A.; Danciu, G. K.; Landemore, H.; Marwick, A.; Miller, C.; Ovadya, A.; Saltz, E.; Schirch, L.; Shalom, D.; Siddarth, D.; Sieker, F.; Small, C.; Stray, J.; Tang, A.; Tessler, M. H.; and Zhang, A. 2024. AI and the future of digital public squares, ArXiv, https://arxiv.org/abs/2412.09988.

Hunter, A., and Potyka, N. 2023. Syntactic reasoning with conditional probabilities in deductive argumentation. *Artificial Intelligence* 321:103934.

Hunter, A., and Thimm, M. 2014. Probabilistic argumentation with epistemic extensions and incomplete information. Technical report, ArXiv.

Hunter, A.; Polberg, S.; Potyka, N.; Rienstra, T.; and Thimm, M. 2021. Probabilistic argumentation: A survey. In Gobbay, D.; Giacomin, M.; Simari, G. R.; and Thimm, M., eds., *Handbook of Formal Argumentation 2*, volume 2. College Publications. 397–444.

Janssen, J.; Vermeir, D.; and De Cock, M. 2008. Fuzzy argumentation frameworks. In *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty IPMU*, 513–520.

Leite, J., and Martins, J. 2011. Social Abstract Argumentation. In *IJCAI International Joint Conference on Artificial Intelligence*, 2287–2292.

Morveli Espinoza, M.; Nieves, J. C.; and Tacla, C. A. 2023a. A gradual semantics with imprecise probabilities for support argumentation frameworks. In *NMR 2023, The 21st International Workshop on Non-Monotonic Reasoning, Rhodes, Greece, September 2-4, 2023*, 84–93. CEUR-WS.

Morveli-Espinoza, M.; Nieves, J. C.; and Tacla, C. A. 2023b. Probabilistic causal bipolar abstract argumentation: An approach based on credal networks. *Annals of Mathematics and Artificial Intelligence* 91(4):517–536.

Potyka, N. 2018. Continuous Dynamical Systems for Weighted Bipolar Argumentation. In *Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona*.

Potyka, N. 2019. Extending modular semantics for bipolar weighted argumentation. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, 1722–1730. Richland, SC: IFAAMAS.

Rago, A., and Toni, F. 2017. Quantitative argumentation debates with votes for opinion polling. In An, B.; Bazzan, A.; Leite, J.; Villata, S.; and van der Torre, L., eds., *PRIMA 2017: Principles and Practice of Multi-Agent Systems*, 369–385. Cham: Springer.

Rago, A.; Toni, F.; Aurisicchio, M.; Baroni, P.; et al. 2016. Discontinuity-free decision support with quantitative argumentation debates. *Fifteenth International Conference on Principles of Knowledge Representation and Reasoning* 16:63–73.

Rossie, J.; Delobelle, J.; Konieczny, S.; Lens, C.; and Vesic, S. 2024. Collective satisfaction semantics for opinion based argumentation. In Marquis, P.; Ortiz, M.; and Pagnucco, M., eds., *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024, Hanoi, Vietnam. November 2-8, 2024*.

Shafer, G. 1976. *A Mathematical Theory of Evidence*, volume 42. Princeton University Press.

Shortall, R.; Itten, A.; Meer, M. v. d.; Murukannaiah, P.; and Jonker, C. 2022. Reason against the machine? future directions for mass online deliberation. *Frontiers in Political Science* Volume 4 - 2022.

Strat, T. M. 1984. Continuous belief functions for evidential reasoning. In *AAAI*, 308–313.

Thieyre, J.; Al Anaissy, C.; Beynier, A.; Destercke, S.; Maudet, N.; and Vesic, S. 2025. Uncertainty in Quantitative Bipolar Argumentation Frameworks. In *Proceedings of the 28th European Conference on Artificial Intelligence (ECAI)*, 1679–1686. IOS Press.

Thimm, M.; Baroni, P.; Giacomin, M.; and Vicig, P. 2018. Probabilities on Extensions in Abstract Argumentation. In Black, E.; Modgil, S.; and Oren, N., eds., *Theory and Applications of Formal Argumentation*, 102–119. Cham: Springer International Publishing.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

Yu, L.; Al Anaissy, C.; Vesic, S.; Li, X.; and van der Torre, L. 2023. A principle-based analysis of bipolar argumentation semantics. In *European Conference on Logics in Artificial Intelligence*, 209–224. Springer.