Depth-Bounded Approximations
of Probability

Paolo Baldi$^{(\boxtimes)}$, Marcello D'Agostino , and Hykel Hosni

Department of Philosophy, University of Milan, Milan, Italy
{paolo.baldi,marcello.dagostino,hykel.hosni}@unimi.it

Abstract. We introduce measures of uncertainty that are based on *Depth-Bounded Logics* [4] and resemble belief functions. We show that our measures can be seen as approximation of classical probability measures over classical logic, and that a variant of the PSAT [10] problem for them is solvable in polynomial time.

1 Introduction

In this work, we investigate the relation between belief functions (see the original [13], and, for a more recent survey, [5]) and probability, from a new logical perspective. Expanding on ideas first introduced in [1], we investigate measures of uncertainty which resemble Dempster-Shafer Belief Functions, but that, instead of being based on classical logic, are based on *Depth-Bounded Logics* (DB logics), a family of propositional logics approximating classical logic [4].

Our starting point is the observation that Belief Functions and Depth-Bounded Logics share a similar concern for the way *virtual* and *actual* information possessed by an agent is evaluated and manipulated.

Let us recall that belief functions can be uniquely determined from so-called mass functions (see e.g. [11]), i.e. probability distributions over the power sets of classical propositional evaluations. If such mass functions are non-zero only for singletons of evaluations, one obtains probability functions, as special cases. We will look at the mass functions behind the probability measures, as arising from the general mass functions (determining arbitrary belief functions) via a limiting process: agents originally assign masses to arbitrary sets of evaluations, reflecting their *actual information*, and they stepwise distribute such mass, only when requested to do so, by way of weighting additional *virtual information*, until they will have their say on the specific uncertainty associated with each single evaluation.

A related issue has been investigated in logic, where the family of DB logics [3,4] relies on the idea of separating two kinds of (classically valid) inferences: the inferences which only serve the purpose to make explicit the information

that agents already possess, i.e. those using only their actual information on the one hand, and those which make use of virtual information on the other. The latter type of inferences arises from the use of a single branching rule (see Fig. 1), reflecting the *principle of bivalence*, which allow agents to reason by cases, adding information not actually in their possession, and drawing further inferences thereon (see Sect. 2). The family of depth-bounded logics is then defined just by fixing maximal depths at which the application of such branching rule is allowed. Unbounded use of the rule results in (an alternative presentation of) classical logic, which can be thus seen as a limit of such family of weaker DB logics.

As an important consequence of the bounded use of the bivalence principle, it is shown in [4] that the consequence relation determined by each DB logic is decidable in polynomial time, hence we can realistically expect that (boundedly) rational agents would be able to recognize, in practice and not only in principle, whether a depth-bounded inference is actually correct. This contrasts with classical logic, which can be seen as the, computationally unfeasible, limit of the feasible DB logics.

The main contribution of this paper is twofold: first, we show that the measures of belief that we introduce, based on DB logics, provide approximations of classical probability measures over classical logic. Second, we prove that under certain reasonable conditions, the problem of finding whether there is any such measure satisfying a given set of linear constraints is solvable in polynomial time, in contrast with the analogous problem for classical logic and probability.

The rest of the paper is structured as follows. In Sect. 2 we recall some preliminaries about DB logics. In Sect. 3 we introduce our depth-bounded measure of uncertainty, based on DB logics, and in Sect. 4 we investigate computational issues. Section 5 contains conclusions and hints at future work.

2   Preliminaries

Let us fix a language $\mathcal{L}$, over a finite set $Var_{\mathcal{L}} = \{p_1, \ldots, p_n\}$ of propositional variables. We let $Fm_{\mathcal{L}}$ be the formulas built from the propositional variables by the usual classical connectives $\wedge, \vee, \neg$, and a constant $\curlywedge$ denoting contradiction. For each $p_i \in Var$ we denote by $\pm p_i$ any of the literals p_i and $\neg p_i$. For each set of formulas Γ we denote by $Sf(\Gamma)$ the subformulas of the formulas in Γ, and by $Var(\Gamma)$ the propositional variables occurring in Γ. Finally by $At_{\mathcal{L}} = \{\pm p_1 \wedge \pm p_2 \wedge \cdots \wedge \pm p_n \mid p_i \in Var_{\mathcal{L}}\}$ we denote the *atoms*, i.e. all the conjunctions of literals, formed from choosing (under the given order) exactly one literal for each of the (finitely many) variables of the language.

Let us now move to consider the family of DB Logics. We start from the 0-depth logic, that is the logic manipulating only actual information. Here we will limit ourselves to a proof-theoretic presentation, based on the Intelim (introduction and elimination) rules in Table 1. For a semantic characterization, see the nondeterministic truth tables, e.g. in [3,4].

The intelim rules determine a notion of 0-depth consequence relation $\vdash_0$, in the usual way.

Table 1. Introduction and elimination rules

$$\frac{\varphi \quad \psi}{\varphi \wedge \psi} \ (\wedge\mathcal{I}) \qquad\qquad \frac{\neg\varphi}{\neg(\varphi \wedge \psi)} \ (\neg\wedge\mathcal{I}1) \qquad \frac{\neg\psi}{\neg(\varphi \wedge \psi)} \ (\neg\wedge\mathcal{I}2)$$

$$\frac{\neg\varphi \quad \neg\psi}{\neg(\varphi \vee \psi)} \ (\neg\vee\mathcal{I}) \qquad\qquad \frac{\varphi}{\varphi \vee \psi} \ (\vee\mathcal{I}1) \qquad\qquad \frac{\psi}{\varphi \vee \psi} \ (\vee\mathcal{I}2)$$

$$\frac{\varphi \quad \neg\varphi}{\curlywedge} \ (\curlywedge\mathcal{I}) \qquad\qquad \frac{\varphi}{\neg\neg\varphi} \ (\neg\neg\mathcal{I})$$

$$\frac{\varphi \vee \psi \quad \neg\varphi}{\psi} \ (\vee\mathcal{E}1) \qquad \frac{\varphi \vee \psi \quad \neg\psi}{\varphi} \ (\vee\mathcal{E}2) \qquad \frac{\neg(\varphi \vee \psi)}{\neg\varphi} \ (\neg\vee\mathcal{E}1)$$

$$\frac{\neg(\varphi \vee \psi)}{\neg\psi} \ (\neg\vee\mathcal{E}2) \qquad \frac{\varphi \wedge \psi}{\varphi} \ (\wedge\mathcal{E}1) \qquad\qquad \frac{\varphi \wedge \psi}{\psi} \ (\wedge\mathcal{E}2)$$

$$\frac{\neg(\varphi \wedge \psi) \quad \varphi}{\neg\psi} \ (\neg\wedge\mathcal{E}1) \qquad \frac{\neg(\varphi \wedge \psi) \quad \psi}{\neg\varphi} \ (\neg\wedge\mathcal{E}2)$$

$$\frac{\neg\neg\varphi}{\varphi} \ (\neg\neg\mathcal{E}) \qquad\qquad \frac{\curlywedge}{\varphi} \ (\curlywedge\mathcal{E})$$

Definition 1. *For any set of formulas $\Gamma \cup \{\alpha\} \subseteq Fm_{\mathcal{L}}$, we let $\Gamma \vdash_0 \alpha$ iff there is a sequence of formulas $\alpha_1, \ldots, \alpha_m$ such that $\alpha_m = \alpha$ and each formula α_i is either in Γ or obtained by an application of the rules in Table 1 on the formulas α_j with $j < i$.*

The key feature of the consequence relation $\vdash_0$ is that only information actually possessed by an agent is allowed in a "0-depth deduction".

As already recalled in the introduction, the DB Logics for $k > 0$, will be defined via the amount k of virtual information which agents are allowed to use in their deductions. This leads to the recursive definition of the consequence relation $\vdash_k$, for $k > 0$, as follows.

Definition 2. *For each $k > 0$ and set of formulas $\Gamma \cup \{\alpha\} \subseteq Fm_{\mathcal{L}}$, we let $\Gamma \vdash_k \alpha$ iff there is a $\beta \in Sf(\Gamma \cup \{\alpha\})$ such that $\Gamma, \beta \vdash_{k-1} \alpha$ and $\Gamma, \neg\beta \vdash_{k-1} \alpha$.*

In other words, we suppose that β and $\neg\beta$ are pieces of "virtual information" which is not actually possessed by the agent, but which is used to derive α through case-based reasoning. While, according to Definition 1, the consequence $\vdash_0$ amounts to the existence of a suitable *sequence* of formulas, the derivability relation $\vdash_k$ amounts to the existence of a suitable *proof-tree*, where each node is labeled by a formula, which is either an assumption or obtained by formulas
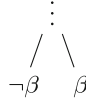
Fig. 1. The branching rule PB (principle of bivalence)

above it by means of an intelim rule, or of the branching rule (PB) in Fig. 1. The latter is then only allowed in a limited form: for $\vdash_k$ we are allowed at most k nested applications of (PB). Thus, $\Gamma \vdash_k \varphi$ can be equivalently taken to say that there is a proof-tree, as described above, so that φ is derivable from Γ in each branch, via the intelim rules, plus the additional virtual information introduced by the branching rules. One may run a proof-search procedure (see e.g. the algorithm in [4]), to verify whether such a proof-tree, deriving φ from Γ in each branch, exists. Even if this is not the case, i.e. if the proof-search procedure only produces proof-trees which derive φ in some (possibly none) and not all of the branches, we are still interested in the structure of such trees, in particular since they keep track of the virtual information that has been explored. This is the main inspiration behind our investigation of depth-bounded belief in the next section.

Before that, let us finish this section recalling two important properties of the DB logics, already mentioned in the introduction, and shown e.g. in [3,4]. First, DB logics provide a hierarchy of consequence relations approximating the classical one, that is, $\vdash_k \subseteq \vdash_{k+1}$ and $\lim_{k \to \infty} \vdash_k = \vdash$, where $\vdash$ stands for classical derivability. Finally, each $\vdash_k$ can be decided in polynomial time, and is thus feasible. This will be of particular use in Sect. 4 of the paper.

3   Depth-Bounded Proofs and Uncertain Reasoning

So far, we have recalled the definition of DB logics and given an idea of how proofs in such logics work, by distinguishing the use of actual and virtual information. Let us now assume that agents, whenever they add a piece of virtual information to their stock of assumptions, can also *weight* their belief on it, for extra-logical reasons. We will then take the belief that an agent commits to a formula φ to be the sum of all the weights assigned to the leaves of a depth-bounded proof-tree, that allow to derive φ. In particular, we request that, if all branches derive φ, which corresponds to φ being logically derivable, one would then obtain a degree of belief 1. We use these ideas as a bridge, from the realm of depth-bounded logic to that of depth-bounded *uncertain reasoning*. Let us recall that classical belief functions can be determined from mass functions, and that, when such mass functions are non-zero only for singletons, one obtain classical probabilities. Identifying formulas and sets of evaluations, one can reformulate this syntactically, by taking the mass functions behind belief functions to act over $Fm_{\mathcal{L}}$, and assume that those behind probabilities are non-zero only over $At_{\mathcal{L}}$ [11]. Our starting point towards depth-bounded uncertain reasoning, is to consider

mass functions which are non-zero only over those formulas which keep track of the information (virtual and actual), used in each branch of a proof-search tree in DB logic. Before delving into our formal definition of depth-bounded belief functions, we will need to fix first various parameters.

- First, we will have a set $In \subseteq Fm_{\mathcal{L}} \cup \{*\}$, the *initial information*, which we assume to be finite. In stands for the formulas, for which an agent, for some extra-logical reasons, can assign a degree of belief, already at a shallow (0-depth) level. We can think of the values of such formulas as obtained from the available data, e.g. as information of statistical nature.

  In order to simplify the notation, we assume that In is nonempty, and we represent the case where no information at all is initially available by the symbol[1] $*$, which is not part of the language, and letting $In = \{*\}$. We adopt the convention that $* \vdash_k \varphi$ stands for $\vdash_k \varphi$.

  In our setting, we can consider a belief conditioned on a formula γ, by just assuming that, for each $\alpha \in In$, we have $\alpha \vdash_0 \gamma$. When this is the case, we say that In is γ-*based* and denote it by In_γ. Similarly we denote the case where $In = \{*\}$ by In_*.

- We have then a set Π, standing for the predictions that an agent wants to obtain, and that thus guide the weighting of her degree of belief. The idea is that an agent weights the uncertainty of virtual information and explores various possible scenarios, only in order to settle, eventually, the truth or falsity of all the formulas in Π.

- Finally, we have a set of virtual information V. This can be thought of as the set of questions that the agent may evaluate, in the process of assessing the formulas in Π. Typical example might be $V = Var(\Pi)$ or $V = Sf(\Pi)$.

Let us recapitulate our setting: starting from initial knowledge in In, agents ask themselves a number of questions about the formulas in V, thus specifying in more details the possible information states, which will be then used to settle the belief and make predictions about the formulas in Π.

We assume that the amount of questions the agents can ask themselves is bounded: the maximum number of questions an agent can ask corresponds, in a sense to be made precise later, to the depth of derivations in DB logic.

We are now ready to give our first formal definition of 0-depth mass functions, representing the initial evidence possessed by an agent. This is nothing else than a convex distribution over the set In of the initial information.

Definition 3. *A* 0-*depth mass function is a function* $m_0 \colon In \to [0,1]$ *such that* $\sum_{\alpha \in In} m_0(\alpha) = 1$ *and* $m_0(\alpha) = 0$ *if* $\alpha \vdash_0 \curlywedge$.

Note that, in case $In = \{\gamma\}$, we have $m_0(\gamma) = 1$ and $m_0(\alpha) = 0$ for any other formula α.

---

[1] We slightly depart from the notation in [4], where the state of no information is denoted by $\perp$, since the latter is often used as a constant for *falsum* in intuitionistic and various nonclassical logics.

Definition 4. *Given a 0-depth mass-function m_0, a 0-depth belief function is a function $B_0\colon Fm_{\mathcal{L}} \to [0,1]$ such that*

$$B_0(\varphi) = \sum_{\substack{\alpha \in In \\ \alpha \vdash_0 \varphi}} m_0(\alpha) \qquad B_0(\varphi) = 0 \ \text{if for no } \alpha \in In, \alpha \vdash_0 \varphi.$$

If In is of the form In_γ for some $\gamma \in Fm_{\mathcal{L}}$, we will then have $B_0(\varphi) = 1$ if $\gamma \vdash_0 \varphi$ and $B_0(\varphi) = 0$ otherwise.

Remark 1. As in the case of classical belief and mass functions, $m_0(\varphi)$ represents a portion of belief committed *exclusively* to φ and to no other formula, while $B_0(\varphi)$ stands for the belief in φ, which is obtained by putting together all the basic pieces of belief leading to (i.e. 0-depth deriving) φ. Note that, while in principle 0-depth equivalent formulas can be assigned different values via a 0-depth mass, they will still be assigned the same 0-depth belief. Hence, in our framework, masses cannot be uniquely determined by belief functions. This is due to fact that we assign masses to formulas, rather than to equivalence classes of the corresponding Lindenbaum-Tarski algebra (see e.g. [11]).

The notions of 0-depth mass and 0-depth belief function encode the shallow information, which is provided to an agent. We will now introduce mass functions based on higher DB logics, corresponding to the setting where agents have both higher inferential and "imaginative" power, i.e. when they can weight the uncertainty of pieces of information going beyond what is originally given.

Let us fix a triplet $\mathcal{G} = \langle In, \Pi, V \rangle$ where $In \subseteq Fm_{\mathcal{L}} \cup \{*\}$, $\Pi \subseteq Fm_{\mathcal{L}}$, $V \subseteq Fm_{\mathcal{L}}$, with the intended meaning discussed above. We will represent the information evaluated by an agent, in the form of forests, with labels provided via the triplet $\mathcal{G}$. Let us recall that by a forest we just mean a disjoint union of trees, in graph-theoretic terms.

Definition 5. *Let F be a binary forest. A $\mathcal{G}$-label for F is a labeling of nodes of F into formulas in $Fm_{\mathcal{L}}$ such that:*

- *When restricted to the roots of the trees in F, the labeling is a bijection with the formulas in In.*
- *For each node labeled by α, the children nodes are labeled by $\alpha \wedge \beta$ and $\alpha \wedge \neg\beta$ for some $\beta \in V$.*

Before proceeding, we also need the following technical definition.

Definition 6. *Let F be any $\mathcal{G}$-labeled forest*

- *We say that a formula γ k-decides δ if $\gamma \vdash_k \delta$ or $\gamma \vdash_k \neg\delta$.*
- *We let $Lf(F)$ be the set of formulas that label the leaves of F.*
- *We say that a leaf labeled by α is Π-closed if $\alpha \vdash_0 \curlywedge$ or α k-decides δ, for each $\delta \in \Pi$. A leaf which is not Π-closed is said to be open.*

We will build now a set of $\mathcal{G}$- labeled forests of a given maximal depth. Each open node is expanded by two new children nodes, representing the addition of a certain piece of virtual information and its negation.

Definition 7. *Let $\mathcal{G} = \langle In, \Pi, V \rangle$. We define recursively the set of $\mathcal{G}$-labeled forests F_k of depth k, for any $k \in \mathbb{N}$, as follows :*

– *For $k = 0$ we let F_0 be a set of nodes with no edges, each labeled by a distinct formula in In. Clearly $Lf(F_0) = In$.*
– *The set F_k, for $k \geq 1$ is the set of all $\mathcal{G}$-labeled forests obtained as follows:*
  • *Pick a $\beta \in V$ and, for each $\mathcal{G}$-labeled forest $F' \in F_{k-1}$, expand each Π-open leaf labeled by α in F' with two nodes labeled by $\alpha \wedge \beta$ and $\alpha \wedge \neg\beta$.*
  • *(MAX) Among the resulting forests, add to F_k only those forests F such that the number of formulas in $Lf(F)$, which 0-depth derive $\pm\varphi$, for each $\varphi \in \Pi$, is maximal[2].*

In the following, for each $F \in F_k$ we call the forest $F' \in F_{k-1}$ from which it was obtained, via the construction above, the predecessor *of F.*

Definition 8. *Let F_k be the set of $\mathcal{G}$-labeled forests of depth k. For each forest $F \in F_k$, we let $m_k^F : Lf(F) \to [0,1]$ be any function such that:*

(i) *$m_k^F(\gamma \wedge \alpha) + m_k^F(\gamma \wedge \neg\alpha) = m_{k-1}^{F'}(\gamma)$ where $F' \in F_{k-1}$ is the predecessor of F, $\gamma \in Lf(F')$ and γ labels the parent node in F of $\gamma \wedge \alpha$ and $\gamma \wedge \neg\alpha$.*
(ii) *$m_k^F(\gamma) = m_{k-1}^{F'}(\gamma)$ if $F' \in F_{k-1}$ is the predecessor of F and $\gamma \in Lf(F') \cap Lf(F)$.*
(iii) *$m_k^F(\gamma) = m_k^G(\delta)$ for each $F, G \in F_k$, $\gamma \in Lf(F), \delta \in Lf(G)$ such that $\gamma \vdash_0 \delta$ and $\delta \vdash_0 \gamma$.*

Recalling Definition 3 and condition (i) in Definition 8, it is easy to see that, for each $F \in F_k$

$$\sum_{\alpha \in Lf(F)} m_k^F(\alpha) = 1$$

Each m_k^F is thus a mass functions, in the sense of Shafer's belief function [13], which is non-zero only over the leaves of the trees in F.

Definition 9. *Let $\mathcal{G} = (In, \Pi, V)$ and F_k be a $\mathcal{G}$-labeled set of forests. For each $F \in F_k$, we define the F-based k-depth belief function B_k^F and the k-depth plausibility function Pl_k^F as follows:*

$$B_k^F(\varphi) = \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \varphi}} m_k^F(\alpha) \qquad Pl_k^F(\varphi) = \sum_{\substack{\alpha \in Lf(F) \\ \alpha \nvdash_0 \neg\varphi}} m_k^F(\alpha)$$

---

[2] This condition might not seem intuitive, but actually plays an important conceptual role, given the motivations of our model. While we want to depart from unrealistic assumptions behind both classical inferences and probability, we still want our models to be *prescriptive*, rather than purely descriptive. In other words, we want to model how agents *should* weight their uncertainty, given their limited inferential ability. Therefore, even if it could be the case that agents use the *wrong* piece of virtual information (i.e. failing the condition (MAX)) we limit ourselves to the case where they only use the virtual information actually leading them to settle as many of their questions as possible, within their inferential abilities.

Finally, we define the $\mathcal{G}$-based k-depth belief as a function B_k from the formulas in $Fm_{\mathcal{L}}$ to the interval subsets of $[0, 1]$, associating to any formula φ the following interval:

$$B_k(\varphi) = [\min_{F \in F_k} B_k^F(\varphi), \max_{F \in F_k} B_k^F(\varphi)]$$

Remark 2. It is immediate to see that $Pl_k^F(\varphi) = 1 - B_k^F(\neg\varphi)$, hence, for each forest F we can think that an exact measure of uncertainty of the formula φ lies within the interval $[B_k^F(\varphi), Pl_k^F(\varphi)]$. Any such interval is related with a single forest F. This should not be confused with the interval given by $[\min_{F \in F_k} B_k^F(\varphi), \max_{F \in F_k} B_k^F(\varphi)]$ which arises from considering various k-depth belief functions over different forests, that is, various proof-search strategies, involving different pieces of virtual information.

We now show some properties of our construction, which highlight its connection with belief functions on the one hand, and with DB logics on the other.

Proposition 1. *(a) Assume $\mathcal{G} = (In_\gamma, \{\varphi\}, Sf(In \cup \{\varphi\}))$, for some $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$, $\varphi \in Fm_{\mathcal{L}}$, and let F_k be the set of $\mathcal{G}$-labeled k-depth forests. If $\gamma \vdash_k \varphi$, then for all the $F \in F_k$, we have $B_k^F(\varphi) = 1$.*

(b) Assume $\mathcal{G} = (In_\gamma, \{\varphi\}, Sf(In \cup \{\varphi\})$ for some $\gamma \in Fm_{\mathcal{L}} \cup \{*\}$, $\varphi \in Fm_{\mathcal{L}}$, and let F_k be the set of $\mathcal{G}$-labeled k-depth forests. If $\gamma \vdash_k \neg\varphi$, then for all $F \in F_k$, we have $B_k^F(\varphi) = 0$.

(c) Assume $\mathcal{G} = (In, \{\varphi, \psi\}, Sf(In \cup \{\varphi, \psi\}))$ for some $\varphi, \psi \in Fm_{\mathcal{L}}$, and let F_k be the set of $\mathcal{G}$-labeled k-depth forests. If $\varphi \vdash_k \psi$, we have:
  - *There is an $F \in F_k$ such that $B_k^F(\varphi) \leq B_k^F(\psi)$*
  - *There is an $l \geq k$ such that, for any forest $F \in F_l$, we get $B_l^F(\varphi) \leq B_l^F(\psi)$.*

(d) Assume $\mathcal{G} = (In, \Pi, V)$, and let F_k be the set of $\mathcal{G}$-labeled k-depth forests. For each $F \in F_k, \varphi_1, \ldots, \varphi_n \in Fm_{\mathcal{L}}$, we have:

$$B_k^F(\bigvee_{i=1}^n \varphi_i) \geq \sum_{\emptyset \neq S \subseteq 1,\ldots,n} (-1)^{|S|-1} B_k^F(\bigwedge_{i \in S} \varphi_i).$$

Proof. (a). Consider the forest G, obtained by attaching to any $\alpha \in In$, the tree containing the virtual information in a k-depth proof of φ from γ. Now, since $\alpha \vdash_0 \gamma$, and each $\beta \in Lf(G)$ contains all the virtual information in a k-depth proof of φ from γ, we will have that $\beta \vdash_0 \varphi$, for all $\beta \in Lf(G)$. By condition (MAX) since there is a forest, G, such that all its leaves derive φ, then all the forests $F \in F_k$ need to have the same property. Hence, we obtain that for each $F \in F_k$, $B_k^F(\varphi) = \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \varphi}} m_k^F(\alpha) = 1.$

(b). By (a), for any forest F, we have $B_k^F(\neg\varphi) = 1$. This means that, for any $\alpha \in Lf(F)$ such that $m_k^F(\alpha) > 0$, $\alpha \vdash_0 \neg\varphi$. If $\alpha \vdash \varphi$ we would get $\alpha \vdash \curlywedge$, which by definition of m_k^F implies $m_k^F(\alpha) = 0$ in contradiction with our assumption. Hence, for any $\alpha \in Lf(F)$, $\alpha \nvdash_0 \varphi$, and $B_k^F(\varphi) = 0$.

(c). The first claim holds, by just taking the forest F to be constituted of the virtual information used in a k-depth proof of ψ from φ. Then, any formula labeling a leaf 0-depth deriving φ, will derive ψ as well. For the second claim, take any forest $F \in F_k$, and consider all the leaves which are 0-depth deriving φ, but not deriving ψ. Expand such leaves with the virtual information contained in any k-depth proof of ψ from φ. This results in a forest of depth l for some $l \geq k$, where each leaf 0-depth deriving φ derives ψ as well. By the maximality condition (MAX) in Definition 7, all forests at depths l will have this property, since otherwise they cannot decide all the formulas in $\Pi = \{\varphi, \psi\}$.

(d). We straightforwardly adapt Theorem 4.1 in [11]. Pick a forest $F \in F_k$. First, for each $\alpha \in Lf(F)$ let $Ind(\alpha) = \{i \mid \alpha \vdash_k \varphi_i\}$. Note that one can show by induction on $|Ind(\alpha)|$, that if $Ind(\alpha) \neq \emptyset$, then $\sum_{\emptyset \neq S, \, S \subseteq Ind(\alpha)} (-1)^{|S|-1} = 1$. We thus get:

$$
\begin{aligned}
B_k^F(\bigvee_{i=1}^n \varphi_i) &= \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 (\varphi_1 \vee \cdots \vee \varphi_n)}} m_k^F(\alpha) \\
&\geq \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \varphi_1 \text{ or } \ldots \text{ or } \alpha \vdash_0 \varphi_n}} m_k^F(\alpha) \\
&= \sum_{Ind(\alpha) \neq \emptyset} m_k^F(\alpha) = \sum_{Ind(\alpha) \neq \emptyset} m_k^F(\alpha) \sum_{\emptyset \neq S \subseteq Ind(\alpha)} (-1)^{|S|-1} \\
&= \sum_{\emptyset \neq S \subseteq Ind(\alpha)} (-1)^{|S|-1} \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \bigwedge_{i \in S} \varphi_i}} m_k^F(\alpha) \\
&= \sum_{\emptyset \neq S \subseteq \{1, \ldots, n\}} (-1)^{|S|-1} B_k(\bigwedge_{i \in S} \varphi_i).
\end{aligned}
$$

Let us now discuss some examples.

Example 1. Let $\mathcal{G} = (\{*\}, \{\alpha \vee \beta\}, \{\alpha, \beta\})$.

At depth 0, we only have the tree with a single node labeled by $*$. We obtain $B_0(\alpha \vee \beta) = 0$ since $* \nvdash_0 \alpha \vee \beta$. At depth 1, our possible forests are actually just trees. Two trees satisfy the constraints in Definition 7, namely :

$$
\begin{array}{cc}
* & * \\
\overparen{\alpha \quad \neg\alpha} & \overparen{\beta \quad \neg\beta}
\end{array}
$$

Let us call the left tree above S and the right one T, and let $m_1^S(\alpha) = 0.5$ and $m_1^S(\neg\alpha) = 0.5$ while $m_1^T(\beta) = 0.4$ and $m_1^T(\neg\beta) = 0.6$. Applying Definition 9, we thus obtain $B_1^S(\alpha \vee \beta) = m_1^S(\alpha) = 0.5$ and $B_1^T(\alpha \vee \beta) = m_1^T(\beta) = 0.4$. Hence $B_1(\alpha \vee \beta) \in [B_1^T(\alpha \vee \beta), B_1^S(\alpha \vee \beta)] = [m_1^T(\beta), m_1^S(\alpha)] = [0.4, 0.5]$. Note that, on the other hand, $B_1(\alpha) \in [B_1^T(\alpha), B_1^S(\alpha)] = [m_1^T(\alpha), m_1^S(\alpha)] = [0, 0.5]$ and $B_1(\beta) \in [B_1^S(\beta), B_1^T(\beta)] = [m_0^S(\beta), m_0^T(\beta)] = [0, 0.4]$. Let us move now to depth

2. In S we only need to expand the node $\neg\alpha$, since α is $\{\alpha \vee \beta\}$-closed (it is sufficient to 0-depth derive $\alpha \vee \beta$). The same holds for $\neg\beta$ in the tree T. We get:

$$
\begin{array}{cc}
* & * \\
\overset{\frown}{\alpha \quad \neg\alpha} & \overset{\frown}{\beta \quad \neg\beta} \\
\overset{\frown}{\beta \quad \neg\beta} & \overset{\frown}{\alpha \quad \neg\alpha}
\end{array}
$$

where for simplicity, we display only the piece of virtual information added by each node, rather than their actual labels (which can be read off by the conjunction of the formula displayed in the node with all its ancestors). Let us call the two trees above again S and T for simplicity. By condition (ii) in Definition 8, we have $m_2^S(\alpha) = m_1^S(\alpha) = 0.5$, and $m_2^T(\beta) = m_1^T(\beta) = 0.4$. For the remaining nodes of S, we let $m_2^S(\neg\alpha \wedge \beta) = 0.2$ and $m_2^S(\neg\alpha \wedge \neg\beta) = 0.3$, and for T, we let $m_2^T(\neg\beta \wedge \alpha) = 0.3$ and $m_2^T(\neg\beta \wedge \neg\alpha) = 0.3$. We obtain finally $B_2^T(\alpha \vee \beta) = m_2^T(\beta) + m_2^T(\neg\beta \wedge \alpha)$ and $B_2^S(\alpha \vee \beta) = (m_2^S(\alpha) + m_2^S(\neg\alpha \wedge \beta)]$, hence

$$B_2(\alpha \vee \beta) \in [B_2^T(\alpha \vee \beta), B_2^S(\alpha \vee \beta)] = [0.7, 0.7]$$

Remark 3. The example above can be generalized considering, for any $n \geq 2$, the triplet $\mathcal{G} = (\{*\}, \{\varphi_1 \vee \cdots \vee \varphi_n\}, \{\varphi_1, \ldots, \varphi_n\})$. The corresponding $\mathcal{G}$- based belief function determines, for each $F \in F_n$ a corresponding permutation σ such that:

$$
\begin{aligned}
B_n^F(\varphi_1 \vee \cdots \vee \varphi_n) = {} & B_1^F(\varphi_{\sigma(1)}) + B_2^F(\neg\varphi_{\sigma(1)} \wedge \varphi_{\sigma(2)}) + \ldots \\
& + B_n^F(\neg\varphi_{\sigma(1)} \wedge \cdots \wedge \neg\varphi_{\sigma(n-1)} \wedge \varphi_{\sigma(n)})
\end{aligned}
$$

Example 2. Let us now consider the famous example of the Ellsberg urn [6]. We assume to have a language with propositional variables {Y,R,B} which stand for the proposition *the next extracted ball is Yellow— Red— Blue*, respectively. The initial knowledge is that 2/3 of the balls are either yellow or red and 1/3 are blue. The background theory is given by the conjunction γ of the formulas in the set

$$\{Y \to (\neg B \wedge \neg R), R \to (\neg B \wedge \neg Y), B \to (\neg R \wedge \neg Y)\}$$

which encode the information that any extracted ball has exactly one of the colors Y, B, R. We now consider the $\mathcal{G}$- based k-depth belief, with $\mathcal{G} = \langle In_\gamma, \Pi, V \rangle$, where

$$In_\gamma = \{(Y \vee R) \wedge \gamma, B \wedge \gamma\} \quad \Pi = \{Y, R, B\} \quad V = \{Y, R\}$$

We formalize the factual information about the proportion of the balls, together with the background theory, by letting: $m_0((Y \vee R) \wedge \gamma) = 2/3, m_0(B \wedge \gamma) = 1/3$.

This implies $B_0(Y \vee R) = 2/3, B_0(B) = 1/3, B_0(\gamma) = 1, B_0(Y) = B_0(R) = 0$. At depth 1, we are required to make use of virtual information. One can easily check that, via the node labeled $B \wedge \gamma$, we can already prove $B \wedge \gamma \vdash_0 \neg Y$, that $B \wedge \gamma \vdash_0 \neg R$ and that $B \wedge \gamma \vdash_0 B$. The node is thus Π-closed, and it should not be expanded. On the other hand, we will need to expand the node $(Y \vee R) \wedge \gamma$ with either the virtual information on Y or on R. We obtain thus a forest $F \in F_1$ of the form:

$$(Y \vee R) \wedge \gamma \qquad\qquad B \wedge \gamma$$

$$\widehat{Y \ \neg Y}$$

and a forest $G \in F_1$ of the form:

$$(Y \vee R) \wedge \gamma \qquad\qquad B \wedge \gamma$$

$$\widehat{R \ \neg R}$$

which have exactly the same structure. So at depth 1, the agent will assign $m_1^F(((Y \vee R) \wedge \gamma) \wedge Y)$ and $m_1^F(((Y \vee R) \wedge \gamma) \wedge \neg Y)$ such that their sum equals $m_0^F((Y \vee R) \wedge \gamma))$. Note that we easily obtain $((Y \vee R) \wedge \gamma) \wedge \neg Y \vdash_0 ((Y \vee R) \wedge \gamma) \wedge R$ and $((Y \vee R) \wedge \gamma) \wedge Y \vdash_0 ((Y \vee R) \wedge \gamma) \wedge \neg R$. The converse direction of the consequence holds as well, hence by condition (iii) in Definition 8, we have: $m_1^G(((Y \vee R) \wedge \gamma) \wedge R) = m_1^G(((Y \vee R) \wedge \gamma) \wedge \neg Y)$ and $m_1^G(((Y \vee R) \wedge \gamma) \wedge \neg R) = m_1^F(((Y \vee R) \wedge \gamma) \wedge Y)$.

At depth 1, considering that the information about the colors is completely symmetric, a natural assumption is now to adopt a uniform distribution, i.e. $m_1((Y \vee R) \wedge Y)) = m_1((Y \vee R) \wedge \neg Y)) = 1/3$. This means that $B_1(Y) = B_1(R) = B_1(B) = 1/3$.

To conclude this section, we now show that we can see usual classical probability functions as arising from sequences of depth-bounded belief functions. By classical probability in our setting, we just mean finitely additive measures, defined as functions $P \colon Fm_{\mathcal{L}} \to [0, 1]$.

Theorem 1. *Let $P \colon Fm_{\mathcal{L}} \to [0, 1]$ be a classical probability function. Then there is a sequence of $\mathcal{G}$-labeled depth bounded belief functions such that, for each formula φ, we have $P(\varphi) = \lim_{k \to \infty} B_k(\varphi)$.*

Proof. Let $\mathcal{G} = (\{*\}, Fm_{\mathcal{L}}, Var_{\mathcal{L}})$. Recalling that $Var_{\mathcal{L}} = \{p_1, \ldots, p_n\}$, we obtain that, for each forest $F \in F_n$, the set $Lf(F)$ coincides with $At_{\mathcal{L}}$, up to permutations of the literals in each atom. Let us consider the mass function m_n over the set $Lf(F)$ such that $m_n(\alpha) = P(\alpha)$ for each $\alpha \in Lf(F)$. Now, we obtain that for any $F \in F_n$

$$P(\varphi) = \sum_{\substack{\alpha \in At_{\mathcal{L}} \\ \alpha \vdash \varphi}} P(\alpha) = \sum_{\substack{\alpha \in At_{\mathcal{L}} \\ \alpha \vdash_0 \varphi}} P(\alpha) = \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \varphi}} m_n(\alpha) = B_n^F(\varphi)$$

All forests in F_n will have the same leaves, modulo a permutations of the literals appearing in the conjunction. Hence, by condition (iii) in Definition 8, for any $F, G \in F_k$, $\alpha \in Lf(F)$ and $\sigma(\alpha) \in G$, where σ is a permutation of the literals in α, we need to have $m_n^G(\sigma(\alpha)) = m_n^F(\alpha) = P(\alpha)$. Hence, by Definition 9, we have $B_n^F(\varphi) = B_n^G(\varphi) = P(\varphi)$ for each $\varphi \in Fm_{\mathcal{L}}$. This implies that the interval for $B_n(\varphi)$ in Definition 9 reduces to the single value $P(\varphi)$. On the other hand, all the leaves in the forests in F_n are $Fm_{\mathcal{L}}$-closed, since atoms decide all the formulas in $Fm_{\mathcal{L}}$, hence $F_k = F_n$ for any $k \geq n$, and $B_k(\varphi) = B_n(\varphi) = P(\varphi)$ for any $k \geq n$. From this the main claim immediately follows.

4   Complexity of Depth-Bounded Belief

In this section we investigate the conditions under which our approach provides a feasible model of reasoning under uncertainty. For concepts in complexity theory, we refer the reader e.g. to [14]. Following previous works based on classical probability, e.g. [7,8,10,11], we assume that an agent is provided n linear constraints over her belief on the formulas $\varphi_1, \ldots, \varphi_m$, of the form:

$$\sum_{j=1}^{m} a_{ij} B(\varphi_j) = w_i \quad i = 1, \ldots, n \quad a_{ij}, w_i \in \mathbb{Q}. \tag{1}$$

Our setup suggests then the following decision problem, which stands to our k-depth logic and k-depth belief functions as the GENPSAT problem (see e.g. [2]) stands to classical logic and classical probability functions:

GEN-B$_0$-SAT Problem

INPUT: The set of m formulas and n linear constraints in (1).

PROBLEM: Is there a 0-depth belief function B_0 over $In = \{\varphi_1, \ldots, \varphi_m\}$ satisfying the n constraints in (1)?

Recalling Definition 9, the problem boils down to finding a solution for the following system of linear inequalities in the unknowns $m_0(\varphi_1), \ldots, m_0(\varphi_m)$.

$$\sum_{j=1}^{m} a_{ij} \sum_{\substack{k=1,\ldots,m \\ \varphi_k \vdash_0 \varphi_j}} m_0(\varphi_k) = w_i \qquad \text{for each } i = 1, \ldots, n$$

$$m_0(\varphi_j) \geq 0 \qquad \text{for each } j = 1, \ldots, m$$

$$\sum_{j=1}^{m} m_0(\varphi_j) = 1$$

$$m_0(\varphi_j) = 0 \qquad \text{if } \varphi_j \vdash_0 \curlywedge$$

Let us denote by $size(In)$ the number of symbols occurring in the formulas in In, and by $inc(In)$ the number of inconsistent formulas among those in In. We recall from [4] that both, finding out whether $\varphi_j \vdash_0 \varphi_i$, and whether $\varphi_j \vdash_0 \curlywedge$ requires time polynomial in $size(In)$. On the other hand, the system above has size $((n + m + 1 + inc(In)) \times m)$, and finding a solution is polynomial as well. Hence the problem above turns out to be in **PTIME**$(size(In) + n)$.

Let us now consider the problem of finding out whether there is a k-depth belief function, for a given $k > 0$, satisfying the constraints in (1). Recalling the Definition 9, this problem amounts to solving a linear system as the one above, where the set In is replaced by the set $Lf(F)$, for all the various $F \in F_k$. Recall that the latter are determined by the parameters Π and V, discussed in the previous section. Let us still set $In = \{\varphi_1, \ldots, \varphi_m\}$, which is given as input to the problem, as the information initially provided to an agent. For

the remaining two parameters, we take $V = Var(In)$ and $\Pi = f(In)$, for some computable function f. We fix a $k > 0$ and consider then the following:

GEN–B$_k$-SAT Problem

INPUT: A triplet $\mathcal{G} = (\{\varphi_1, \ldots, \varphi_m\}, f(\{\varphi_1 \ldots \varphi_m\}), Var(In))$ and the n constraints in (1).

PROBLEM: Is there a $\mathcal{G}$-based k-depth belief function B_k^F, with F in the $\mathcal{G}$-labeled set of forests F_k, which satisfies the n constraints in (1)?

Answering to this problem corresponds to finding an F in F_k, for which the following system, in the unknowns $m_k^F(\alpha)$, has a solution:

$$\sum_{j=1}^{m} a_{ij} \sum_{\substack{\alpha \in Lf(F) \\ \alpha \vdash_0 \varphi_j}} m_k^F(\alpha) = w_i \qquad \text{for each } i = 1, \ldots, n$$

$$m_k^F(\alpha) \geq 0 \qquad \text{for each } \alpha \in Lf(F)$$

$$\sum_{\alpha \in Lf(F)} m_k^F(\alpha) = 1$$

$$m_k^F(\alpha) = 0 \qquad \text{if } \alpha \vdash_0 \curlywedge$$

This problem also turns out to be polynomial, if the size of Π is polynomially bounded. We give a sketch of proof in the following.

Theorem 2. GEN-B$_k$-SAT *can be decided in* **PTIME**$(size(In) + n + size(\Pi))$.

Proof. By our construction, for any forest $F \in F_k$ the number of leaves in $Lf(F)$ is bounded above by $|In| \cdot 2^k$, which is linear in $|In|$ (once k is fixed, 2^k is constant). The number of possible forests, on the other hand, is bounded by the number of subsets of $|Var(In)|$ of cardinality k, which is polynomial in $|Var(In)|$ and, consequently, polynomial in $size(In)$. Indeed, we can safely disregard any permutation or repetitions of the same virtual information, due to condition (iii) of Definition 8 and condition (MAX) of Definition 7, respectively. We then need to do some "pruning" among all possible forests, by discarding the branches which are not Π-open and the forest that do not satisfy the maximality condition (MAX) in Definition 7. The latter is obtained then, by running, whenever necessary, for each formula in Π and its negation, the polynomial time algorithm, e.g. in [4]. Once we have determined the set F_k, we have then, for each $F \in F_k$ a set of formulas in $Lf(F)$. We have then to look for a solution to the system above. Each such system has size $(n+|Lf(F)|+1+inc(Lf(F)) \times |Lf(F)|$, hence it is still polynomially bounded. Finally, since solving each system requires polynomial time, we obtain the claim.

Finally, let us notice that, if we take $Size(\Pi) = f(Size(In))$, where f is a polynomially bounded function, then the GEN-B$_k$-SAT Problem is in **PTIME**$(Size(In) + n)$. As an example, this would hold if we take, as a reasonable choice $\Pi = Sf(In)$.

5   Conclusions and Future Work

In this work we have introduced feasible approximations of probability measures, based on Depth-Bounded logics. The resulting measures shed light on the connection between two approximation problems: the approximation of probability, as a limiting case of belief functions and that of classical logic as a limiting case of depth-bounded boolean logic. In future research, we plan to compare our approach with the Transferable Belief Model of [15], and similar works, which handle the relation between belief functions and probability. While the former are considered in [15] to be adequate to model the *credal*, i.e. purely mental, aspect of belief, the latter are taken as good models for its *pignistic* aspect, i.e. its role as a guide towards decisions. Decision-theoretic models are also a natural setting to evaluate and deepen our results. In particular, in the context of subjective expected utility, various weakenings of Savage axioms [12] have been considered in the literature (see e.g. [9] for an overview). We plan to investigate how these works relate to our approach, which weakens instead the logic.

References

1. Baldi, P., Hosni, H.: Depth-bounded belief functions. Int. J. Approximate Reasoning (2020). https://doi.org/10.1016/j.ijar.2020.05.001
2. Caleiro, C., Casal, F., Mordido, A.: Generalized probabilistic satisfiability. Inf. Comput. **332**, 39–56 (2017)
3. D'Agostino, M.: An informational view of classical logic. Theor. Comput. Sci. **606**, 79–97 (2015)
4. D'Agostino, M., Finger, M., Gabbay, D.M.: Semantics and proof-theory of depth-bounded boolean logics. Theor. Comput. Sci. **480**, 43–68 (2013)
5. Denoeux, T.: 40 years of Dempster-Shafer Theory. Int. J. Approximate Reasoning **79**, 1–6 (2016)
6. Ellsberg, D.: Risk, Ambiguity, and the Savage Axioms. Quart. J. Econ. **75**(4), 643–669 (1961)
7. Fagin, R., Halpern, J.Y., Megiddo, N.: A logic for reasoning about probabilities. Inf. Comput. **1−2**(87), 78–128 (1990)
8. Finger, M., De Bona, G.: Probabilistic satisfiability: algorithms with the presence and absence of a phase transition. Ann. Math. Artif. Intell. **75**(3–4), 351–389 (2015)
9. Gilboa, I., Marinacci, M.: Ambiguity and the Bayesian Paradigm, pp. 385–439. Springer, Cham (2016)
10. Hansen, P., Jaumard, B.: Probabilistic satisfiability. In: Kohlas, J., Moral, S. (eds.) Handbook of Defeasible Reasoning and Uncertainty Management Systems. Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 5, pp. 321–367. Springer, Kluwer, Dordrecht (2000). https://doi.org/10.1007/978-94-017-1737-3_8
11. Paris, J.B.: The Uncertain Reasoner's Companion: A Mathematical Perspective. Cambridge University Press, Cambridge (1994)
12. Savage, L.J.: The Foundations of Statistics. 2nd edn., Dover (1972)

13. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
14. Sipser, M.: Introduction to the Theory of Computation. PWS Publishing Company (1997)
15. Smets, P., Kruse, R.: The Transferable Belief Model for Belief Representation. In: Motro, A., Smets, P. (eds.) Uncertainty Management in Information Systems, pp. 343–368. Springer, Boston (1997). https://doi.org/10.1007/978-1-4615-6245-0_12