Revision Operators with Compact Representations*

Pavlos Peppas[a], Mary-Anne Williams[b], Grigoris Antoniou[c,d]

[a] *University of Patras, pavlos@upatras.gr*

[b] *University of New South Wales, mary-anne.williams@unsw.edu.au*

[c] *University of Huddersfield, G.Antoniou@hud.ac.uk*

[d] *L3S, Leibniz University Hannover, Germany*

Abstract

Despite the great theoretical advancements in the area of Belief Revision, there has been limited success in terms of implementations. One of the hurdles in implementing revision operators is that their specification (let alone their computation), requires substantial resources. On the other hand, implementing a specific revision operator, like Dalal's operator, would be of limited use. In this paper we generalise Dalal's construction, defining a whole family of concrete revision operators, called *Parametrised Difference revision operators* or *PD operators* for short. This family is wide enough to cover a wide range of different applications, and at the same time it is easy to represent. In addition to its semantic definition, we characterise the family of PD operators axiomatically (including a characterisation specifically for Dalal's operator), we prove its' compliance with Parikh's relevance-sensitive postulate (P), we study its computational complexity, and discuss its benefits for belief revision implementations.

Keywords: Belief Revision

1. Introduction

The AGM framework [3] is the dominant paradigm for the study of *belief revision*. It has been studied extensively and lies on solid theoretical foundations (see [4] for a

*This submission is the journal version (with improvements and additions) of the research work previously published in conference papers; in particular in [1] and (partially) in [2].

survey). Yet despite the success of its theoretical models, little has been done in terms of implementations of AGM belief revision operators. This is not to say that important attempts have not been made; see for example [5], [6], [7]. None of them however has had the great impact on real-world applications that one would expect from a successful belief revision solver.

There are at least two major obstacles to a successful implementation of an AGM belief revision system that can work beyond toy examples. The first is the high computational complexity of the belief revision process [8]; we will have more to say about this later in the paper.

The second is the large amount of information that, in principle, the user needs to provide explicitly to the system; we shall call this the *representational cost* of a belief revision solver. Recall that the AGM postulates for revision specify not one, but an entire class of revision functions. Hence, before a belief revision system can answer any query about the result of revising a theory K by a sentence φ, denoted $K * \varphi$, the user needs to specify the particular revision function $*$ under consideration. There are many ways that this can be done, but in principle, they are all equivalent to specifying a family of total preorders over possible worlds; i.e., one total preorder for each theory of the object language L [9]. For a propositional language with n variables, there exist 2^{2^n} theories; clearly an enormous number. Even if one focuses only on a single theory K, one still needs to specify a preorder $\preccurlyeq$ over the 2^n possible worlds.

Of course there are shortcuts. For example one can request only a *partial* specification of a preorder over worlds and fill in the remaining information automatically, using some (intuitive) default rule. In this case it is important that the side-effects of the completion process are well understood, and that the formal properties of the resulting revision functions are thoroughly investigated. The other option is to avoid the requirement for preorder specification altogether by choosing to implement only one *concrete* revision operator. The problem of course is that such a system would be rather limited in scope. Moreover, as far as concrete "off-the-shelve" AGM revision operators go, there are not all that many to choose from. Out of the few well known proposals, like [10], [11], [12], [13], it is only Dalal's operator [14] that satisfies all the AGM postulates for revision.

Herein we introduce an entire class of concrete revision operators, all of which satisfy the full set of AGM postulates for revision; we call them *Parametrised Difference revision operators*, or *PD operators* for short. PD operators are essentially generalisations of Dalal's operator. Most importantly, each PD operator can be fully specified from a preorder over the *n propositional variables* (also called *atoms*) of the object language L. In other words, a *single* preorder over the n atoms suffices to generate the preorders over possible worlds associated with *all* 2^{2^n} theories of L. This is a *double exponential* drop on the representational cost.

The range of applicability of PD operators is illustrated via some characteristic belief revision scenarios, including ones from iterated revision, that are beyond the reach of Dalal's operator.

Perhaps surprisingly, this added expressiveness of PD operators compared to Dalal's operator comes at no extra computational cost: we show that PD operators lie at the same level of the polynomial hierarchy as Dalal's operator. Moreover, and probably more importantly for practical applications, when confined to Horn formulas, and the size of the new evidence is small compared to that of the initial knowledge base, PD operators become tractable; in fact, complexity can be further reduced to *linear time* with respect to the size of the initial knowledge base, if the the size of the queries is bounded by a constant.

PD operators are defined semantically and characterised axiomatically via six conditions named (D1) – (D6). As a by-product we also provide an axiomatic characterisation of Dalal's operator. Moreover we illustrate some attractive formal properties of PD operators including their compliance with Parikh's notion of relevance-sensitive revision [15].

The rest of the paper is structured as follows. The next section introduces some notation and terminology, followed by a section covering the necessary background on AGM belief revision. In Section 4 we introduce the class of PD operators and illustrate their use in characteristic belief revision scenarios. Then, new axioms characterising

PD operators are formulated, accompanied by corresponding representation results. The formal properties of PD operators are discussed in Section 6. The following section discusses the representational cost of AGM revision, and compares the effectiveness of PD operators with previous approaches in dealing with this problem. This is followed by a section that contains our study on the computational complexity of PD operators. In the last section we provide some concluding remarks.

2. Preliminaries

In this article we shall be working with a propositional language L built over *finitely many* propositional variables. The finite, nonempty set of all propositional variables (also called *atoms*) is denoted by P. A *literal* is a variable in P or the negation of a variable. If l is a literal containing the variable α, then by $\bar{l}$ we denote the literal $\neg\alpha$ if $l = \alpha$, and the literal α otherwise. The letters x, y, p, q, and z (possibly with subscripts and/or superscripts) will always represent literals. The letters A, B, C, D, and E (possible with subscripts and/or superscripts) will always represent sets of literals. For a set of literals A, we define $\overline{A}$ to be the set $\overline{A} = \{\overline{q} : q \in A\}$. We will sometimes abuse notation and treat a set of literals A as a sentence, namely the conjunction of all its literals, leaving it to the context to resolve any ambiguity; thus for example, in "$A \subseteq B$", A is a set of literals whereas in "$\neg A$", A is a sentence. In the limiting case where the *set* A is empty, we take the *sentence* A to be an arbitrary tautology.

An *interpretation* assigns truth values to propositional variables; more formally, an interpretation v (over P) is a function mapping every propositional variable in P to the set $\{T, F\}$, where "T" stands for "true" and "F" stands for "false". The definition of v can be extended to assign truth values to arbitrary sentences of L using the classical semantics of the Boolean connectives: $v(\neg\varphi) = T$ iff $v(\varphi) = F$, $v(\varphi \vee \psi) = T$ iff $v(\varphi) = T$ or $v(\psi) = T$, $v(\varphi \wedge \psi) = T$ iff $v(\varphi) = T$ and $v(\psi) = T$, $v(\varphi \rightarrow \psi) = v(\neg\varphi \vee \psi)$, and $v(\varphi \leftrightarrow \psi) = v((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$. For any sentence $\varphi \in L$, we shall say that v *satisfies* φ, which we denote as $v \models \varphi$, iff $v(\varphi) = T$. We shall say that v satisfies a set of sentences Γ, denoted $v \models \Gamma$, iff v satisfies all sentences in Γ.

Finally, a sentence or set of sentences is called *consistent* iff there exists at least one interpretation that satisfies it.

The set of all interpretations over P is denoted $\mathcal{M}$. Interpretations will also be called *possible worlds* (or simply *worlds*) and henceforth will be identified with the set of literals they satisfy. Like with any set of literals, a possible world r may sometimes be treated as a sentence (as for example in the expression "$r \vee r'$)". Possible worlds will be denoted with the letters w, r, u, possibly with subscripts and/or superscripts. Arbitrary sentences of L will be denoted by the Greek letters φ, ψ (possibly with subscripts and/or superscripts).

For a set of sentences Γ and a sentence φ of L, we shall write $\Gamma \models \varphi$ iff every interpretation that satisfies Γ, also satisfies φ. For sentence $\varphi, \psi \in L$, we shall write $\psi \models \varphi$ as an abbreviation of $\{\psi\} \models \varphi$. For sets of sentences Γ, Δ, we shall write $\Gamma \models \Delta$ iff $\Gamma \models \varphi$, for all $\varphi \in \Delta$.

For a set of sentences Γ of L, by $Cn(\Gamma)$ we denote the set of all logical consequences of Γ, i.e., $Cn(\Gamma) = \{\varphi \in L \colon \Gamma \models \varphi\}$. A theory K of L is any set of sentences of L closed under $\models$, i.e., $K = Cn(K)$. We shall use the letters K, H, and T to denote theories of L. The set of all consistent theories is denoted by $\mathcal{K}$. A theory K is complete iff for all sentences $\varphi \in L$, $\varphi \in K$ or $\neg\varphi \in K$.

For a set of sentences Γ of L, $[\Gamma]$ denotes the set of all possible worlds that satisfy Γ. Often we shall use the notation $[\varphi]$ for a sentence $\varphi \in L$, as an abbreviation of $[\{\varphi\}]$. For a theory K and a set of sentences Γ of L, we shall denote by $K + \Gamma$ the closure under $\models$ of $K \cup \Gamma$, i.e., $K + \Gamma = Cn(K \cup \Gamma)$. For a sentence $\varphi \in L$ we shall often write $K + \varphi$ as an abbreviation of $K + \{\varphi\}$.[1]

[1]In this paper we essentially follow the notation and terminology that is typically used in the Belief Revision literature.

3. The AGM Framework

In the AGM framework, [3], [16], belief revision is modelled as a function $*$ mapping a theory K and a sentence φ, to a theory $K*\varphi$, representing the result of revising K by φ. Alchourrón, Gärdenfors and Makinson have introduced a set of eight postulates, numbered below as $(K*1)$ – $(K*8)$, that ought to be satisfied by any *rational* revision function. These postulates are now known as the *AGM postulates for revision*, and the functions that satisfy these postulates are known as *AGM revision functions* (or simply *revision functions*):[2]

$(K*1)$ $K*\varphi$ is a theory of L.

$(K*2)$ $\varphi \in K*\varphi$.

$(K*3)$ $K*\varphi \subseteq K+\varphi$.

$(K*4)$ If $\neg\varphi \notin K$ then $K+\varphi \subseteq K*\varphi$.

$(K*5)$ If φ is consistent then $K*\varphi$ is also consistent.

$(K*6)$ If $\models \varphi \leftrightarrow \psi$ then $K*\varphi = K*\psi$.

$(K*7)$ $K*(\varphi \wedge \psi) \subseteq (K*\varphi)+\psi$.

$(K*8)$ If $\neg\psi \notin K*\varphi$ then $(K*\varphi)+\psi \subseteq K*(\varphi \wedge \psi)$.

It turns out that any AGM revision function can be constructed with the use of a set of total preorders over possible worlds [9]; one total preorder $\preccurlyeq_K$ for each theory K. Recall that a total preorder $\preccurlyeq_K$ over $\mathcal{M}$ is any binary relation over $\mathcal{M}$ that is reflexive and transitive, and such that for all $w, w' \in \mathcal{M}$, $w \preccurlyeq_K w'$ or $w' \preccurlyeq_K w$. As usual, $\prec_K$ denotes the strict part of $\preccurlyeq_K$. Moreover, we shall write $w \approx_K w'$ iff $w \preccurlyeq_K w'$ and $w' \preccurlyeq_K w$.

[2]For a detailed discussion of the AGM postulates see [16] or [4].

A total preorder $\preccurlyeq_K$ is said to be *faithful to* K iff for all $w, w' \in \mathcal{M}$, (i) if $w \in [K]$ then $w \preccurlyeq_K w'$, and, (ii) if $w \in [K]$ and $w' \notin [K]$ then $w \prec_K w'$.[3].

Given a faithful preorder $\preccurlyeq_K$ for each theory K, one can construct a revision function $*$ by means of the following condition, [9]:[4]

$$(\preccurlyeq *) \quad [K * \varphi] = \min([\varphi], \preccurlyeq_K).$$

In the above definition, $min(S, \preccurlyeq_K)$ is the set of *minimal* elements of the set S with respect to $\preccurlyeq_K$; i.e., $min(S, \preccurlyeq_K) = \{w \in S$: for all $w' \in S$, if $w' \preccurlyeq_K w$, then $w \preccurlyeq_K w'\}$. Hence according to $(\preccurlyeq *)$, $K * \varphi$ is defined as the theory satisfied precisely by the $\preccurlyeq_K$-minimal worlds in $[\varphi]$.

Katsuno and Mendelzon have shown that the functions induced from total faithful preorders via $(\preccurlyeq *)$ are exactly those satisfying the AGM postulates for revision. Moreover, since we assume herein that the set P of propositional variables in *finite*, it holds that for any given AGM revision function $*$ and theory K, there is a *unique* total preorders $\preccurlyeq_K$, that satisfies $(\preccurlyeq *)$. We shall call this unique total preorder, *the faithful preorder that* $*$ *assigns to* K.

For ease of presentation, in the rest of the paper we shall focus only on revision of *consistent* theories by *consistent* sentences. Hence from now on, unless explicitly stated otherwise, we assume that the initial belief set K is a consistent theory, and that the epistemic input φ is a consistent sentence.

4. Parametrised Difference Operators

As evident from the previous section, to fully describe an AGM revision function $*$ one needs to specify the faithful preorder $\preccurlyeq_K$ assigned to each theory K of L. For a

[3]It should be noted that in [9], Katsuno and Mendelzon define faithfulness for the broader class of *partial preorders*; wherein however, it suffices to confine ourselves to total preorders.

[4]We note that in fact Katsuno and Mendelzon assign faithful preorders to *sentences* rather than to *theories*. However this difference is inconsequential under the assumptions of this paper.

propositional language built over n atoms, there exist 2^n possible worlds, and $2^{2^n} - 1$ consistent theories. A general belief revision solver that requires users to provide a preorder over 2^n worlds for each consistent theory of L (or even for one theory of L), will not be of any use beyond toy examples.

One can of course trade generality for representational efficiency and build a specialised belief revision solver; one that implements a single concrete AGM revision operator $*$. In this case the faithful preorders that define $*$ can be "hard-wired" into the solver. Users won't need to provide any background information – just the knowledge base and the epistemic input.

Perhaps the most popular concrete AGM revision operator is Dalal's operator [14], and its computational complexity has been studied in Eiter and Gottlob's seminal article [8].

Dalal provides a very natural way of defining the preorder $\preccurlyeq_K$ associated to a theory K. We note that $\preccurlyeq_K$ is meant to encode the comparative plausibility of possible worlds: the closer a world is to the beginning of the preorder the more plausible it is. Dalal defines plausibility in terms of a notion of *difference* between worlds.

In particular, for any two worlds $w, r \in \mathcal{M}$, the difference between w and r, denoted $\mathit{Diff}(w, r)$, is defined to be the set of propositional variables over which the two worlds disagree; i.e., $\mathit{Diff}(w, r) = \{q \in P\colon w \models q \text{ and } r \models \overline{q}\} \cup \{q \in P\colon r \models q \text{ and } w \models \overline{q}\}$. The preorder $\sqsubseteq_K$ that Dalal assigns to a consistent theory K is defined as follows: for all $r, r' \in \mathcal{M}$, $r \sqsubseteq_K r'$ iff there is a $w \in [K]$ such that for all $w' \in [K]$, $|\mathit{Diff}(w, r)| \leqslant |\mathit{Diff}(w', r')|$. Dalal's operator, which we denote $\square$, is defined as the revision function induced from $\{\sqsubseteq_K\}_{K \in \mathcal{K}}$.

An example of Dalal's preorder for a language L built from only three variables a, b, c, assigned to the theory $K = Cn(\{a, b, c\})$, is given below:[5]

[5]We note that $\sqsubset_K$ denotes the *strict* part of $\sqsubseteq_K$; i.e. $r \sqsubset_K r'$ iff $r \sqsubseteq_K r'$ and $r' \not\sqsubseteq_K r$.

$$
\begin{array}{ccccccc}
 & & ab\overline{c} & & a\overline{b}\overline{c} & & \\
abc & \sqsubset_K & a\overline{b}c & \sqsubset_K & \overline{a}b\overline{c} & \sqsubset_K & \overline{a}\overline{b}\overline{c} \\
 & & \overline{a}bc & & \overline{a}\overline{b}c & &
\end{array}
$$

In the above example, the plausibility of a world r is determined by the number of propositional variables over which r differs from the initial world abc.

However, a belief revision solver restricted only to Dalal's operator would be of limited practical use. Many interesting belief revision scenarios lie outside the scope of Dalal's operator (see below).

Clearly we need to strike the right balance between expressivity and representational cost. Our answer to this is the class on PD operators introduced below.

A silent assumption in Dalal's approach is that *all atoms have the same epistemic value*; hence for example, a change in atom a is assumed to be as plausible (or implausible) as a change in atom b. This is clearly a severe restriction that limits considerably the range of applicability of Dalal's operator. PD operators lift this restriction, allowing propositional variables to have different epistemic values.

Suppose for example that for a certain application, the atoms a and b have greater epistemic value than the atom c, and consequently a change in a or b is less plausible than a change in c. This can be encoded by a total preorder $\trianglelefteq$ over the variables a, b, c as follows: $c \vartriangleleft a$, $c \vartriangleleft b$, $a \trianglelefteq b$, and $b \trianglelefteq a$.[6] Given $\trianglelefteq$ we can *refine* Dalal's preoder to take into account the difference in epistemic value between a, b, and c:

$$
abc \sqsubset_K^{\trianglelefteq} ab\overline{c} \sqsubset_K^{\trianglelefteq} \begin{array}{c} a\overline{b}c \\ \\ \overline{a}bc \end{array} \sqsubset_K^{\trianglelefteq} \begin{array}{c} a\overline{b}\overline{c} \\ \\ \overline{a}b\overline{c} \end{array} \sqsubset_K^{\trianglelefteq} \overline{a}\overline{b}c \sqsubset_K^{\trianglelefteq} \overline{a}\overline{b}\overline{c}
$$

In the example above the ranking of possible worlds takes place in two stages. The first stage is identical to Dalal's ranking: each world r is ranked according to the num-

[6]As usual $\vartriangleleft$ denotes the strict part of $\trianglelefteq$; moreover the obvious relationships $a \trianglelefteq a$, $b \trianglelefteq b$, and $c \trianglelefteq c$ have been omitted for the sake of readability.

ber of switches in propositional variables that are necessary to turn the initial world abc into r. At the second stage the ranking is further refined to take into account the different epistemic value of the propositional variables that have been switched. In particular, for any two worlds r, r' that require the same number of switches from abc (i.e., $|\mathit{Diff}(abc, r)| = |\mathit{Diff}(abc, r')|$), r is more plausible than r' iff $\mathit{Diff}(abc, r)$ *lexicographically precedes* $\mathit{Diff}(abc, r')$ with respect to $\trianglelefteq$. Thus for example, $ab\overline{c} \sqsubset_K^{\trianglelefteq} \overline{a}bc$ because $c \lhd a$ (despite the fact that both worlds are one switch away from abc).

The example above illustrates the basic idea in generalising Dalal's approach. The formal definition of PD preorders is given below.

Firstly, some auxiliary notation. For a total preorder $\trianglelefteq$ over P, a set $S \subseteq P$ and an atom $q \in P$, by S_q we denote the set $S_q = \{p \in S : p \trianglelefteq q\}$. The definition of $\trianglelefteq$ can now be extended to *sets* of propositional variables as follows:

Definition 1. *Let $\trianglelefteq$ be a total preorder over P. We define the extension of $\trianglelefteq$ over sets of atoms as follows. For any $S, S' \subseteq P$, $S \trianglelefteq S'$ iff one of the following three conditions holds:*

(a) $|S| < |S'|$.

(b) $|S| = |S'|$, *and for all* $q \in P$, $|S_q| = |S'_q|$.

(c) $|S| = |S'|$, *and for some* $q \in P$, $|S_q| > |S'_q|$, *and for all* $p \lhd q$, $|S_p| = |S'_p|$.

In the above definition, condition (b) states that S and S' are lexicographically indistinguishable with respect to $\trianglelefteq$, whereas (c) states that S lexicographically precedes S' (wrt $\trianglelefteq$). Henceforth we shall use $\trianglelefteq$ to denote both a total preorder over atoms, as well as its extension over sets of atoms; any ambiguity will be resolved by the context.

It is not hard to verify that (the extended) $\trianglelefteq$ is a total preorder over 2^P:

Lemma 1. *Let $\trianglelefteq$ be any total preorder over the atoms in P. The extension of $\trianglelefteq$ over sets of atoms as defined above, is a total preorder over 2^P.*

Proof. Reflexivity follows immediately from the definition of $\trianglelefteq$ (in particular, condition (b)). For totality, consider any two sets of atoms S, S'. If $|S| < |S'|$ then by condition (a) of the above definition, $S \trianglelefteq S'$; likewise, if $|S'| < |S|$ then $S' \trianglelefteq S$ and thus, once again, S and S' comparable with respect to $\trianglelefteq$. Assume therefore that $|S| = |S'|$. If for all $q \in P$, $|S_q| = |S'_q|$, then by condition (b), $S \trianglelefteq S'$. If not, there is a $\trianglelefteq$-minimal atom in $q \in P$ such that $|S_q| \neq |S'_q|$. Without loss of generality we can assume that $|S_q| > |S'_q|$. Observe that since q is a $\trianglelefteq$-minimal atom for which $|S_q| \neq |S'_q|$, it follows that $|S_p| = |S'_p|$ for all $p \lhd q$. Hence by condition (c), $S \trianglelefteq S'$; i.e. S and S' are again $\trianglelefteq$-comparable. Thus $\trianglelefteq$ is total.

For transitivity, consider any three sets of atoms S, S', S'', such that $S \trianglelefteq S'$ and $S' \trianglelefteq S''$. Assume towards contradiction that $S \ntrianglelefteq S''$. Since, as shown above, $\trianglelefteq$ is total, we derive that $S'' \trianglelefteq S$. Hence, by the definition of $\trianglelefteq$, $|S''| \leqslant |S|$. Moreover, from $S \trianglelefteq S'$ and $S' \trianglelefteq S''$ we derive respectively that $|S| \leqslant |S'|$ and $|S'| \leqslant |S''|$. Thus it follows that $|S| = |S'| = |S''|$. Then from $S'' \lhd S$ we derive that there is a $q \in P$ such that $|S''_q| > |S_q|$ and $|S''_p| = |S_p|$ for all $p \lhd q$.

Observe that if $|S''_p| = |S'_p|$ for all $p \in P$, then $|S'_q| > |S_q|$ and $|S'_p| = |S_p|$ for all $p \lhd q$; this contradicts $S \trianglelefteq S'$. Hence from $S' \trianglelefteq S''$ we derive that there is a $z \in P$ such that $|S'_z| > |S''_z|$, and $|S'_p| = |S''_p|$ for all $p \lhd z$. Next we distinguish between two cases, and we show that in both we reach a contradiction. Firstly assume that $q \lhd z$. Then $|S'_q| = |S''_q| > |S_q|$, and for all $p \lhd q$, $|S'_p| = |S''_p| = |S_p|$. This entails $S \ntrianglelefteq S'$, which of course contradicts our initial assumption. Assume on the other hand that $z \trianglelefteq q$. Then, $|S'_z| > |S''_z| \geqslant |S_z|$. Moreover, for all $p \lhd z$, $|S'_p| = |S''_p| = |S_p|$. Therefore, once again, $S \ntrianglelefteq S'$; contradiction. $\square$

The intended reading of $\trianglelefteq$, defined over sets of variables, is the same as before: $S \trianglelefteq S'$ means that S' as a whole is at least as important as S (as a whole). Therefore, if during believe revision there was a choice between changing all variables in S or changing all variables in S', we would pick the former.

Based on this reading, we define the *PD preorder* $\sqsubseteq_K^{\trianglelefteq}$ *over* $\mathcal{M}$*, induced from* $\trianglelefteq$ *at a theory* K, as follows:

(AW) $r \sqsubseteq_K^{\trianglelefteq} r'$ iff there is a $w \in [K]$ such that for all $w' \in [K]$, $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w', r')$.

Lemma 2. *Let $\trianglelefteq$ be any total preorder over the atoms in P and K a consistent theory of L. The binary relation $\sqsubseteq_K^{\trianglelefteq}$ over $\mathcal{M}$ induced from $\trianglelefteq$ at K via (AW), is a total preorder over $\mathcal{M}$, and it is faithful to K.*

Proof. Let r, r' be any two worlds in $\mathcal{M}$. Moreover, let w_r and $w_{r'}$ be two worlds in $[K]$ such that $\mathit{Diff}(w_r, r)$ is minimal with respect to (the extended) $\trianglelefteq$ in the set $\{\mathit{Diff}(u, r) : u \in [K]\}$, and, likewise, $\mathit{Diff}(w_{r'}, r')$ is minimal with respect to $\trianglelefteq$ in the set $\{\mathit{Diff}(u', r') : u' \in [K]\}$. Since, by Lemma 1, $\trianglelefteq$ is a *total* preorder, it follows from (AW) that $r \sqsubseteq_K^{\trianglelefteq} r'$ iff $\mathit{Diff}(w_r, r) \trianglelefteq \mathit{Diff}(w_{r'}, r')$.

It is now quite straightforward to prove the lemma. Since $\trianglelefteq$ is transitive, reflexive, and total, then so is $\sqsubseteq_K^{\trianglelefteq}$.

For faithfulness, assume that $r \in [K]$. Then, clearly, $\mathit{Diff}(w_r, r) = \mathit{Diff}(r, r) = \varnothing$. Moreover from the definition of $\trianglelefteq$, it follows immediately that $\varnothing \trianglelefteq S$ for any set of atoms S; hence $\mathit{Diff}(w_r, r) \sqsubseteq_K^{\trianglelefteq} \mathit{Diff}(w_{r'}, r')$ for any world r'. Thus, all worlds in $[K]$ are minimal with respect to $\sqsubseteq_K^{\trianglelefteq}$. For the converse, let r' we any world not in $[K]$. Then clearly, $\mathit{Diff}(w_{r'}, r') \neq \varnothing$. On the other hand, for any world $r \in [K]$, $\mathit{Diff}(w_r, r) = \varnothing$. Thus, $|\mathit{Diff}(w_r, r)| < |\mathit{Diff}(w_{r'}, r')|$ and consequently, from the definition of $\trianglelefteq$ we derive that $\mathit{Diff}(w_r, r) \triangleleft \mathit{Diff}(w_{r'}, r')$. This again entails $r \sqsubset_K^{\trianglelefteq} r'$, and therefore r' is not minimal in $\mathcal{M}$ with respect to $\sqsubseteq_K^{\trianglelefteq}$. Hence $\sqsubseteq_K^{\trianglelefteq}$ is faithful to K. $\square$

Notice that according to this definition, a *single* preorder $\trianglelefteq$ over P suffices to determine the preorders assigned to *all* consistent theories K. Hence a preorder $\trianglelefteq$ generates a family of PD preorders $\{\sqsubseteq_K^{\trianglelefteq}\}_{K \in \mathcal{K}}$ which in turn define a revision function $*$. A revision function so constructed is called a *Parametrised Difference revision operator* or a *PD operator* for short.

To illustrate the use of PD operators in encoding belief revision scenarios, consider the following modified version of an example in [17]. A circuit consists of two adders and one multiplier. The variables a_1, a_2, and m represent the facts that "adder1 is working", "adder2 is working", and "the multiplier is working" respectively. Initially we believe that the circuit is working properly. Moreover we know that multipliers are less reliable than adders. Hence, if we observe that there is a malfunction in the circuit, it is plausible to assume that the multiplier (rather than one of the adders) is not working properly.

This scenario can easily be encoded with a PD operator. In particular, consider the PD operator $*$ induced from the following preorder $\trianglelefteq$ on the propositional variables a_1, a_2, m: $m \lhd a_1$, $m \lhd a_2$, $a_1 \trianglelefteq a_2$, and $a_2 \trianglelefteq a_1$ (in addition, $\trianglelefteq$ includes all pairs that follow from reflexivity and transitivity). It is not hard to verify that with this preorder, the revision of $Cn(\{a_1, a_2, m\})$ by $\neg a_1 \vee \neg a_2 \vee \neg m$ leads us to $Cn(\{a_1, a_2, \neg m\})$ as desired.[7]

Our next example comes from [18]: "We encounter a strange new animal and it appears to be a bird, so we believe the animal is a bird. As it comes closer to our hiding place, we see clearly that the animal is red, so we believe that it is a red bird. To remove further doubts about the animal's birdhood, we call in a bird expert who takes it for examination and concludes that it is not really a bird but some sort of mammal. The question now is whether we should still believe that the animal is red."

Once again PD operators deliver the anticipated results. Let us denote by a the proposition "the animal is red" and by b the proposition "the animal is a bird". Our initial belief set is $Cn(\{b\})$. Let $*$ by any PD operator and $\trianglelefteq$ the preorder over atoms associated with $*$. Since a is consistent with $Cn(\{b\})$ it follows that $Cn(\{b\}) * a = Cn(\{a, b\})$. Let $K = Cn(\{a, b\})$. Clearly, $|\mathit{Diff}(ab, a\overline{b})| < |\mathit{Diff}(ab, \overline{a}\overline{b})|$ and therefore $a\overline{b} \sqsubset_K^{\trianglelefteq} \overline{a}\overline{b}$, regardless of the preorder $\trianglelefteq$. Hence $Cn(\{b\}) * a * \neg b = Cn(\{a, \neg b\})$ as desired.

It should be noted that, not only classical AGM, but even its extension with the postulates for iterated revision proposed by Darwiche and Pearl in [17], has trouble dealing with such examples (see [19], [18], and [20] for details).

[7] On the other hand, Dalal spreads the blame equally to all three components of the circuit; i.e., the Dalal-revision of $Cn(\{a_1, a_2, m\})$ by $\neg a_1 \vee \neg a_2 \vee \neg m$ yields the theory $Cn(\overline{a_1} a_2 m \vee a_1 \overline{a_2} m \vee a_1 a_2 \overline{m})$.

Our last example is new. It is essentially an extension of the curcuit example in [17].

A circuit consists of a multiplier and two adders. Let us denote by m the proposition "the multiplier is working", and by a_1, a_2 the propositions "the first adder is working" and "the second adder is working" respectively. After performing some tests on the circuit, we discover that the multiplier or adder 1 is malfunctioning; in symbols, $\neg m \vee \neg a_1$. Suppose that our initial belief set is $Cn(\{m, a_1, a_2\})$. Since it is known that multipliers are less reliable than adders, we end up with the belief set $Cn(\{\overline{m}, a_1, a_2\})$. For the same reason, if our initial belief set were $Cn(\{m, a_1, \overline{a}_2\})$, then $\neg m \vee \neg a_1$ would have taken us to $Cn(\{\overline{m}, a_1, \overline{a}_2\})$. What would then be our response to $\neg m \vee \neg a_1$ had our initial belief set been $Cn(\{m, a_1\})$? Given our past preference to adder 1 over the multiplier (regardless of the status of adder 2), we argue that it is reasonable to once again put the blame on the multiplier. Moreover, since adder 2 is independent from the other two components, our beliefs about adder 2 should not be affected.

The desired behaviour can be easily captured by a PD operator. Indeed the PD operator $*$ used in the first circuit example can also be used for this one. That is, $*$ is induced from a preorder over atoms $\trianglelefteq$ such that $m \lhd a_1$, $m \lhd a_2$, $a_1 \trianglelefteq a_2$, and $a_2 \trianglelefteq a_1$. Clearly by construction, $*$ has all desired properties: $Cn(\{m, a_1, a_2\}) * (\overline{m} \vee \overline{a}_1) = Cn(\{\overline{m}, a_1, a_2\})$, $Cn(\{m, a_1, \overline{a}_2\}) * (\overline{m} \vee \overline{a}_1) = Cn(\{\overline{m}, a_1, \overline{a}_2\})$, and $Cn(\{m, a_1\}) * (\overline{m} \vee \overline{a}_1) = Cn(\{\overline{m}, a_1\})$.

It should be noted that the idea of generalising Dalal's notion of distance between worlds, by differentiating between atoms has been used in the *Belief Merging* literature for quite some time. In particular, preorders on a *weighted Hamming distance* are quite similar to PD preorders. A weighted Hamming distance assigns a *numerical* value (i.e., a *weight*) to each variable of the language. The distance between two possible worlds is then defined as the sum of the weights of all variables over which the two worlds differ (see for example, [21]). These numerical weights assigned to variables can be thought of as the quantitative analog of the preorder $\trianglelefteq$ over variables used in the construction of a PD preorder. There is a major difference however between preorders induced from weighted Hamming distances and PD preorders: with the former it is possible for three worlds w, r, r', to be such that r' is closer to w than r, even though r differs from w in fewer variables than r' (i.e., $|Diff(w,r)| < |Diff(w,r')|$);[8] this can never be the case with PD preorders.[9]

5. Axioms for PD Operators

In this section we provide an axiomatic characterisation of PD operators.

We note from the outset that our new axioms are *not* on a par with the AGM postulates. In fact the two have a totally different purpose. AGM postulates encode general principles of rational belief change. Our new axioms on the other hand, are simply formal properties that characterise a certain class of AGM revision functions (namely those induced from PD preorders), thus providing insight to their behaviour.

Formulating the new axioms was not trivial. The task is complicated by the fact that a PD operator $*$ is constructed from a preorder $\trianglelefteq$ over P in two stages; first $\trianglelefteq$ induces $\{\sqsubseteq_K^{\trianglelefteq}\}_{K\in\mathcal{K}}$, which in turn induces $*$. Thus, metaphorically speaking, $*$ is *two steps away* from its generator $\trianglelefteq$. That makes it harder to devise constraints on $*$ that would project correctly, at a two-steps distance, to $\trianglelefteq$.

For the sake of readability we shall introduce the new axioms in stages. At each stage we provide representation results that highlight the role of the new axioms in the overall characterisation of PD operators.

We recall that throughout this paper, x, y, p, q, z denote literals, A, B, C, D, E denote nonempty consistent sets of literals, φ, ψ denote consistent sentences, and K, H, T denote consistent theories. Moreover, we shall often use concatenation as an abbreviation for conjunction; thus for example AB is an abbreviation of $A \wedge B$, and Ap is an abbreviation of $A \wedge p$.

---

[8]This can happen for example if one of the variables in $Diff(w,r)$ has a weight that is greater than the sum of all weights in $Diff(w,r')$.

[9]We thank the anonymous reviewer of [1] for pointing out previous work on weighted Hamming distances and the similarity of their induced preorders to PD preorders.

For nonempty sets of literals A, B, we define $A \leqq_K B$ iff $A, B \subseteq K$ and $\neg(\overline{A}) \notin K * (\overline{A} \vee \overline{B})$. Intuitively, $A \leqq_K B$ holds whenever, starting from the belief set K (which contains both A and B), it is at least as costly to change (the values of) all literals in B as it is to change all literals in A. We define $A \prec_K B$ as $A \leqq_K B$ and $B \not\leqq_K A$ (or equivalently, $A, B \subseteq K$ and $\neg\overline{B} \in K * (\overline{A} \vee \overline{B})$). Finally, for literals p, q, we define $p \leqq_K q$ and $p \prec_K q$ to be an abbreviation of $\{p\} \leqq_K \{q\}$ and $\{p\} \prec_K \{q\}$, respectively.

5.1. The Special Case of Consistent Complete Theories

Let us start by assuming that the initial belief set K is a consistent *complete* theory. This assumption will allow us to arrive quickly at preliminary representation results that will be instrumental in establishing the general results of the next subsection.

Our first axiom says that if during revision a choice exists between reversing all literals in A or reversing all literals in B, then revision never picks the larger set; in other words, the more literals one needs to reverse during revision, the more costly it is:

(D1) If K is complete and $A \leqq_K B$, then $|A| \leqslant |B|$.

(D1) alone suffices to characterise an interesting super-class of PD operators. In particular, consider the following constraint on a total faithful preorder $\preccurlyeq_K$ assigned to K:[10]

(H) If K is complete and $|Diff(K, r)| < |Diff(K, r')|$, then $r \prec_K r'$.

We shall call a total preorder $\preccurlyeq_K$ over $\mathcal{M}$ satisfying (H), a *Hamming preorder*.[11]

[10] Recall that in this subsection, K is assumed to be a consistent complete theory. Hence we will often abuse notation, like in $Diff(K, r)$, and identify K with the unique possible world that satisfies it.

[11] We note that Hamming preorders are similar, but not quite the same as the preorders induced from weighted Hamming distances discussed earlier.

Theorem 1. *Let K be a consistent complete theory, $*$ an AGM revision function and $\leqslant_K$ the faithful preorder that $*$ assigns to K. Then $\leqslant_K$ is a Hamming preorder iff $*$ satisfies (D1) at K.*

Proof. Assume that $\leqslant_K$ is a Hamming preorder. We prove (D1) by showing its counter-positive. Consider therefore any two sets of literals A, B in K, such that $|B| < |A|$. Define r to be the world that satisfies $\overline{B}$ and agrees with K on all remaining literals, and let r' be any world that satisfies $\overline{A}$. Clearly $|\mathit{Diff}(K,r)| = |B| < |A| \leqslant |\mathit{Diff}(K,r')|$. Hence by (H), $r \prec_K r'$, and therefore no $\overline{A}$-world is minimal in $[\overline{A} \vee \overline{B}]$. This again entails $\neg\overline{A} \in K * (\overline{A} \vee \overline{B})$, and proves (D1).

For the converse, assume that $*$ satisfies (D1) at K and let r, r' be any two worlds such that $|\mathit{Diff}(K,r)| < |\mathit{Diff}(K,r')|$. Clearly $r' \notin [K]$. Hence, if $\mathit{Diff}(K,r) = \varnothing$, then from the faithfulness of $\leqslant_K$ to K we derive that $r \prec_K r'$ as desired. Assume therefore that $\mathit{Diff}(K,r) \neq \varnothing$. Define A to be the set of literals in K over $\mathit{Diff}(K,r)$ and B the set of literals in K over $\mathit{Diff}(K,r')$.[12] Clearly $|A| < |B|$ and therefore by (D1), $\neg\overline{B} \in K * (\overline{A} \vee \overline{B})$. Hence there is a minimal $\overline{A}$-world, call it z, which is strictly smaller wrt $\leqslant_K$ than every $\overline{B}$-world. Consequently, $z \prec_K r'$. Next we show that $r = z$. Suppose on the contrary that $r \neq z$. Then, by construction, there is a $q \in r \cap K$ such that $\overline{q} \in z$. Define $C = A \cup \{q\}$. Then by (D1), $\neg\overline{C} \in K * (\overline{A} \vee \overline{C})$ and therefore $\overline{A} \in K * (\overline{A} \vee \overline{C})$. Hence, since $C = A \cup \{q\}$, from $\neg\overline{C} \in K * (\overline{A} \vee \overline{C})$ we derive that $q \in K * (\overline{A} \vee \overline{C})$ and consequently $z \notin [K * (\overline{A} \vee \overline{C})]$. This however contradicts our original assumption that z is a minimal $\overline{A}$-world. Thus, $r = z$ and therefore $r \prec_K r'$ as desired. $\square$

All PD preorders are Hamming preorders, but not the other way around. Let us take a closer look at the difference between the two.

Given the initial world K, all remaining worlds can be partitioned according to the number of atoms in which they differ from K. In both PD and Hamming preorders, the worlds that differ from K in one atom, precede those that differ from K in two atoms, which precede those that differ from K in three atoms, etc. On the other hand, the rela-

[12] More precisely, $A = \{x \in K : x \in \mathit{Diff}(K,r)\} \cup \{\neg x \in K : x \in \mathit{Diff}(K,r)\}$. Likewise, for B.

tive order of the worlds that belong to the *same* partition is quite different in Hamming and PD preorders: in Hamming preorders the ordering with a partition is arbitrary, whereas in PD preorders it is highly regulated. More precisely, in a PD preorder, the way that the worlds in the first partition are ordered, *fully determines* the ordering of the worlds in all subsequent partitions. In other words, if two PD preorders, faithful to K, agree on the ordering of worlds that differ from K on one atom, then the two preorders are identical. This observation has been the basis for formulating the extra axioms required for PD revisions:

(D2) If K is complete, $A \leqq_K B$, $p \leqq_K q$, and $q \notin B$, then $Ap \leqq_K Bq$.

Axiom (D2) essentially says that if switching the literals in A is at least as easy as switching the literals in B, and switching p is at least as easy as switching q, then switching A and p together is at least as easy as switching B and q together (provided that q is not already in B).

(D3) If K is complete, $A \leqq_K B$, $p \prec_K q$, and $q \notin B$, then $Ap \prec_K Bq$.

Axiom (D3) is essentially the strict version of (D2). Like in (D2), we assume that reversing A is at least as easy as reversing B, but this time we assume that reversing p is strictly easier than reversing q. In this case, says (D3), reversing A and p together is strictly easier than reversing B and q together (provided that $q \notin B$).

(D4) If K is complete, $A \prec_K B$, $p \in K$, $q \notin B$, and for all $z \in B$, $z \leqq_K q$, then $Ap \prec_K Bq$.

Axiom (D4) is based on a similar intuition as (D2) and (D3), but deals with a different case. Suppose that reversing A is strictly easier than reversing B. Moreover assume that reversing the literal q is at least as hard as reversing any literal z in B. Then, says (D4), for *any* literal $p \in K$, changing A and p together is strictly easier than changing B and q together (provided that q is not already in B).

Theorem 2. *Let K be a consistent complete theory, $*$ an AGM revision function and $\leqslant_K$ the faithful preorder that $*$ assigns to K. If $\leqslant_K$ is a PD preorder then $*$ satisfies (D1) – (D4) at K.*

Proof. Assume that $\leqslant_K$ is a PD preorder. Then there exists a preorder $\trianglelefteq$ over P, such that the preorder $\sqsubseteq_K^{\trianglelefteq}$ generated from $\trianglelefteq$ is identical to $\leqslant_K$.

To proceed with the proof we first need to introduce some more notation. For any variable $q \in P$, by q_K we denote q itself if $q \in K$, and the literal $\neg q$ otherwise. Clearly, since K is complete, $q_K \in K$ for all $q \in P$. For a set of variables $A \subseteq P$, by A_K we denote the set $A_K = \{q_K : q \in A\}$.

Next we show that for all $p, q \in P$, $p \trianglelefteq q$ iff $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$. Consider any $p, q \in P$ such that $p \trianglelefteq q$ and suppose towards contradiction that $p_K \in K * (\overline{p_K} \vee \overline{q_K})$. From the latter we derive that there is a $\overline{q_K}$-world, call it r, such that $r \sqsubset_K^{\trianglelefteq} r'$ for all $r' \in [\overline{p_K}]$. Define r'' to be the world that agrees with K on all literals except q. Then clearly, $\mathit{Diff}(K, r'') = \{q\} \subseteq \mathit{Diff}(K, r)$. Hence $\mathit{Diff}(K, r'') \trianglelefteq \mathit{Diff}(K, r)$ and consequently, $r'' \sqsubseteq_K^{\trianglelefteq} r$. Next define u to be the world that agrees with K on all literals except p. Thus $\mathit{Diff}(K, u) = \{p\}$. Given that $u \models \overline{p_K}$, we derive that $r \sqsubset_K^{\trianglelefteq} u$, and consequently, $r'' \sqsubset_K^{\trianglelefteq} u$. Therefore $\mathit{Diff}(K, r'') \lhd \mathit{Diff}(K, u)$, which leads us to $q \lhd p$ contradicting our initial assumption $p \trianglelefteq q$. Hence we have shown that if $p \trianglelefteq q$ then $p \notin K * (\overline{p_K} \vee \overline{q_K})$.

For the converse, suppose that $p, q \in P$ are such that $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$. Then there is a $\overline{p_K}$-world, call it r, such that $r \sqsubseteq_K^{\trianglelefteq} r'$ for all $r' \in [\overline{q_K}]$. Let r'' be the the world that agrees with K on all literals except p. Then $\mathit{Diff}(K, r'') = \{p\} \subseteq \mathit{Diff}(K, r)$. Consequently, $r'' \sqsubseteq_K^{\trianglelefteq} r$. Define u to be the world that agrees with K on all literals except q. Thus $\mathit{Diff}(K, u) = \{q\}$. Given that $u \models \overline{q_K}$, we derive that $r \sqsubseteq_K^{\trianglelefteq} u$, and consequently, $r'' \sqsubseteq_K^{\trianglelefteq} u$. This again entails that $p \trianglelefteq q$ as desired. Hence we have shown that for all $p, q \in P$, $p \trianglelefteq q$ iff $p \notin K * (\overline{p_K} \vee \overline{q_K})$.

We can now proceed to show the validity of the postulates (D1) – (D4).

For (D1), let $A, B \subseteq P$ be such that $\neg\overline{A_K} \notin K * (\overline{A_K} \vee \overline{B_K})$. We will show that

$|A| \leqslant |B|$. Assume on the contrary that $|B| < |A|$. Call r the world that differs from K only over the variables in B. Then clearly, $\mathit{Diff}(K, r) = B$. Moreover, for any $\overline{A_K}$-world r', $A \subseteq \mathit{Diff}(K, r')$. Therefore, from $|B| < |A|$ we derive that for any $\overline{A_K}$-world r', $|\mathit{Diff}(K, r)| < |\mathit{Diff}(K, r')|$, and consequently, $\mathit{Diff}(K, r) \lhd \mathit{Diff}(K, r')$. This again entails $r \sqsubset_K^{\unlhd} r'$, for all $r' \in [\overline{A_K}]$, and consequently, $\neg\overline{A_K} \in K * (\overline{A_K} \vee \overline{B_K})$. Contradiction.

For (D2), consider any $p, q \in P$ and $A, B \subseteq P$ such that $q \notin B$. Assume that $\neg\overline{A_K} \notin K * (\overline{A_K} \vee \overline{B_K})$ and $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$. We will show that $\neg(\overline{A_K}\,\overline{p_K}) \notin K * (\overline{A_K}\,\overline{p_K} \vee \overline{B_K}\,\overline{q_K})$. From $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$ it follows that $p \unlhd q$. Moreover from $\neg\overline{A_K} \notin K * (\overline{A_K} \vee \overline{B_K})$ we derive that there is a $\overline{A_K}$-world, call it r, such that $r \sqsubseteq_K^{\unlhd} r'$, for all $r' \in [\overline{A_K}] \cup [\overline{B_K}]$. Define r'' to be the world that differs from K only over the variables in A. Clearly then, since $r \in [\overline{A_K}]$, we derive that $\mathit{Diff}(K, r'') = \{A\} \subseteq \mathit{Diff}(K, r)$. Consequently, $r'' \sqsubseteq_K^{\unlhd} r$. Next define u to be the world that differs from K only over the variables in B. Then, $\mathit{Diff}(K, u) = B$ and $u \models \overline{B_K}$. Hence, $r \sqsubseteq_K^{\unlhd} u$, and consequently $r'' \sqsubseteq_K^{\unlhd} u$. This again entails that $A \unlhd B$. From this, $p \unlhd q$, and $q \notin B$, it is not hard to derive that $A \cup \{p\} \unlhd B \cup \{q\}$. Next define w to be the world that differ from K only over the variables in $A \cup \{p\}$. Clearly then, $w \models \overline{A_K}\,\overline{p_K}$ and $\mathit{Diff}(K, w) = A \cup \{p\}$. Moreover observe that for any $\overline{B_K}\,\overline{q_K}$-world w', $B \cup \{q\} \subseteq \mathit{Diff}(K, w')$ and therefore $B \cup \{q\} \unlhd \mathit{Diff}(K, w')$. Since $A \cup \{p\} \unlhd B \cup \{q\}$ we then derive that $w \sqsubseteq_K^{\unlhd} w'$ for all $w' \in [\overline{B_K}\,\overline{q_K}]$. This again entails $\neg(\overline{A_K}\,\overline{p_K}) \notin K * (\overline{A_K}\,\overline{p_K} \vee \overline{B_K}\,\overline{q_K})$ as desired.

For (D3), consider any $p, q \in P$ and $A, B \subseteq P$ such that $q \notin B$. Assume that $\neg\overline{A_K} \notin K * (\overline{A_K} \vee \overline{B_K})$ and $q_K \in K * (\overline{p_K} \vee \overline{q_K})$. We will show that $\neg\overline{B_K}\,\overline{q_K} \in K * (\overline{A_K}\,\overline{p_K} \vee \overline{B_K}\,\overline{q_K})$. Firstly observe that from $q_K \in K * (\overline{p_K} \vee \overline{q_K})$ we derive that $p \lhd q$. Moreover from $\neg\overline{A_K} \notin K * (\overline{A_K} \vee \overline{B_K})$ we derive that there is a $\overline{A_K}$-world, call it r, such that $r \sqsubseteq_K^{\unlhd} r'$, for all $r' \in [\overline{A_K}] \cup [\overline{B_K}]$. Define r'' to be the world that differs from K only over the variables in A. Clearly then, since $r \in [\overline{A_K}]$, it follows that $\mathit{Diff}(K, r'') = \{A\} \subseteq \mathit{Diff}(K, r)$. Consequently, $r'' \sqsubseteq_K^{\unlhd} r$. Next define u to be the world that differs from K only over the variables in B. Then, $\mathit{Diff}(K, u) = B$ and $u \in [\overline{B_K}]$. Hence, $r \sqsubseteq_K^{\unlhd} u$, and consequently $r'' \sqsubseteq_K^{\unlhd} u$. This again entails that $A \unlhd B$, which in turn, when combined with $p \lhd q$, and $q \notin B$, leads to $A \cup \{p\} \lhd B \cup \{q\}$. Next define w to be the world that differ from K only over the variables in $A \cup \{p\}$. Clearly then, $w \in [\overline{A_K}\,\overline{p_K}]$ and $\mathit{Diff}(K, w) = A \cup \{p\}$.

Moreover observe that for any $\overline{B_K}\overline{q_K}$-world w', $B \cup \{q\} \subseteq \mathit{Diff}(K, w')$ and therefore $B \cup \{q\} \trianglelefteq \mathit{Diff}(K, w')$. Since $A \cup \{p\} \lhd B \cup \{q\}$ we then derive that $w \sqsubset_K^{\trianglelefteq} w'$ for all $w' \in [\overline{B_K}\overline{q_K}]$. This again entails $\neg\overline{B_K}\overline{q_K} \in K * (\overline{A_K}\overline{p_K} \vee \overline{B_K}\overline{q_K})$ as desired.

Finally for (D4), consider any $p, q \in P$ and $A, B \subseteq P$ such that $q \notin B$. Assume that $\neg\overline{B_K} \in K * (\overline{A_K} \vee \overline{B_K})$ and for all $z \in B$, $z_K \notin K * (\overline{z_K} \vee \overline{q_K})$. We will show that $\neg\overline{B_K}\overline{q_K} \in K * (\overline{A_K}\overline{p_K} \vee \overline{B_K}\overline{q_K})$. First observe that from $z_K \notin K * (\overline{z_K} \vee \overline{q_K})$ for all $z \in B$, we derive that $z \trianglelefteq q$ for all $z \in B$. In other words, q is $\trianglelefteq$-*maximal* in $B \cup \{q\}$. Moreover from $\neg\overline{B_K} \in K * (\overline{A_K} \vee \overline{B_K})$ we derive that there is a $\overline{A_K}$-world, call it r, such that $r \sqsubset_K^{\trianglelefteq} r'$, for all $r' \in [\overline{B_K}]$. Define r'' to be the world that differs from K only over the variables in A. Clearly then, since $r \in [\overline{A_K}]$, it follows that $\mathit{Diff}(K, r'') = \{A\} \subseteq \mathit{Diff}(K, r)$. Consequently, $r'' \sqsubseteq_K^{\trianglelefteq} r$. Next define u to be the world that differs from K only over the variables in B. Then, $\mathit{Diff}(K, u) = B$ and $u \in [\overline{B_K}]$. Hence, $r \sqsubset_K^{\trianglelefteq} u$, and consequently $r'' \sqsubset_K^{\trianglelefteq} u$. This again entails that $A \lhd B$. Then because q is $\trianglelefteq$-*maximal* in $B \cup \{q\}$, it is not hard to verify that $A \cup \{p\} \lhd B \cup \{q\}$, for any variable $p \in P$. Define w to be the world that differ from K only over the variables in $A \cup \{p\}$. Clearly, $w \in [\overline{A_K}\overline{p_K}]$ and $\mathit{Diff}(K, w) = A \cup \{p\}$. Moreover observe that for any $\overline{B_K}\overline{q_K}$-world w', $B \cup \{q\} \subseteq \mathit{Diff}(K, w')$ and therefore $B \cup \{q\} \trianglelefteq \mathit{Diff}(K, w')$. Since $A \cup \{p\} \lhd B \cup \{q\}$ we then derive that $w \sqsubset_K^{\trianglelefteq} w'$ for all $w' \in [\overline{B_K}\overline{q_K}]$. This again entails $\neg\overline{B_K}\overline{q_K} \in K * (\overline{A_K}\overline{p_K} \vee \overline{B_K}\overline{q_K})$. □

The converse of Theorem 2 is also true. To prove it though we need an auxiliary result.

Consider the condition (DD) below (as usual, p denotes a literal, and A denotes a set of literals):

(DD) If K is complete, $p \in K$, $A \subseteq K$ and $p \notin A$, then $p \in K * \overline{A}$.

Lemma 3. *Let K be a consistent complete theory, and $*$ an AGM revision function. If $*$ satisfies (D1) at K then it satisfies (DD).*

Proof. Assume that K is a complete theory and $*$ an AGM revision function that satisfies (D1) at K. Let $\preccurlyeq_K$ be the faithful preorder that $*$ assigns to K. By Theorem 1,

$\preccurlyeq_K$ satisfies (H) at K. Let $A \subseteq K$ be an arbitrary set of literal in K, and p a literal in K such that $p \notin A$. Let r' be any minimal $\overline{A}$-world with respect to $\preccurlyeq_K$. To prove (DD) it suffices to show that $r' \models p$. Assume on the contrary that $r' \models \neg p$. Let p_v be the propositional variable that appears in p. Clearly, since $p \in K$ and $r' \models \neg p$, it follows that $p_v \in \mathit{Diff}(K, r')$. We define r to be the world that agrees with r' over all variables except p_v. Since $p \notin A$, and $r' \models \overline{A}$, by construction it follows that $r \models \overline{A}$. Moreover, again by construction, $\mathit{Diff}(K, r) = \mathit{Diff}(K, r') - \{p_v\}$, and therefore, $|\mathit{Diff}(K, r)| < |\mathit{Diff}(K, r')|$. Condition (H) then entails that $r \prec_K r'$, which however contradicts our assumption that r' is a minimal $\overline{A}$-world with respect to $\preccurlyeq_K$. $\square$

With Lemma 3 we can now prove the converse of Theorem 2

Theorem 3. *Let K be a consistent complete theory, $*$ an AGM revision function and $\preccurlyeq_K$ the faithful preorder that $*$ assigns to K. If $*$ satisfies (D1) – (D4) at K then $\preccurlyeq_K$ is a PD preorder.*

Proof. Assume that $*$ satisfies (D1) – (D4) at K, and let $\preccurlyeq_K$ be the faithful preorder that $*$ assigns to K. We shall construct a preorder $\trianglelefteq$ over P, such that the induced PD preorder $\sqsubseteq_K^{\trianglelefteq}$ over $\mathcal{M}$, coincides with $\preccurlyeq_K$.

First some notation. For any propositional variable $z \in P$, we define z_K to be z itself if $K \models z$, and $\neg z$ otherwise. Since K is assumed to be compete, it clearly follows that $K \models z_K$ for all $z \in P$.

We construct the preorder $\trianglelefteq$ over P as follows:

$$p \trianglelefteq q \quad \text{iff} \quad p_K \notin K * (\neg p_K \vee \neg q_K)$$

Firstly observe that $\trianglelefteq$ is indeed a total preorder over P. Reflexivity follows immediately from the construction of $\trianglelefteq$, and so does totality.

For transitivity, assume that $p \trianglelefteq q \trianglelefteq z$. From $p \trianglelefteq q$ we derive that $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$, and consequently there exists a $\overline{p_K}$-world, call it r, that is $\preccurlyeq_K$-minimal in $[\overline{p_K} \vee \overline{q_K}]$. Likewise, from $q \trianglelefteq z$ we derive that there exists a $\overline{q_K}$-world, call it r', that is $\leqslant_K$-minimal in $[\overline{q_K} \vee \overline{z_K}]$. Since r is $\leqslant_K$-minimal in $[\overline{p_K} \vee \overline{q_K}]$ and $r' \in [\overline{q_K}]$, it follows that $r \leqslant_K r'$. Moreover, since r' is $\leqslant_K$-minimal in $[\overline{q_K} \vee \overline{z_K}]$ and given that $[\overline{z_K}] \subseteq [\overline{q_K} \vee \overline{z_K}]$, we derive that $r' \leqslant_K r''$ for all $r'' \in [\overline{z_K}]$. From the transitivity of $\leqslant_K$ we then derive that $r \leqslant_K r''$ for all $r'' \in [\overline{z_K}]$. This again entails that r is $\leqslant_K$-minimal in $[\overline{p_K} \vee \overline{z_K}]$, and consequently $p_K \notin K * (\overline{p_K} \vee \overline{z_K})$. Hence, $p \trianglelefteq z$ as desired.

Let $\sqsubseteq_K^{\trianglelefteq}$ be the preorder over $\mathcal{M}$ induced from $\trianglelefteq$ at K. Consider two arbitrary worlds $r, r' \in \mathcal{M}$. To complete the proof it suffices to show that

$$r \sqsubseteq_K^{\trianglelefteq} r' \text{ iff } r \leqslant_K r'$$

Since both $\sqsubseteq_K^{\trianglelefteq}$ and $\leqslant_K$ are reflexive and they are both faithful to K, the above equivalence follows immediately if $r \in [K]$ or $r' \in [K]$ or $r = r'$. Assume therefore that $r, r' \notin [K]$ and $r \neq r'$.

First consider the (easy) case where $|Diff(K, r)| \neq |Diff(K, r')|$. Without loss of generality, let's assume that $|Diff(K, r)| < |Diff(K, r')|$. Then clearly, by definition, $r \sqsubset_K^{\trianglelefteq} r'$. Moreover from Theorem 1, it also follows that $r \prec_K r'$.

Hence, if $|Diff(K, r)| \neq |Diff(K, r')|$, it holds that $r \sqsubseteq_K^{\trianglelefteq} r'$ iff $r \leqslant_K r'$ as desired.

Next we consider the case where $|Diff(K, r)| = |Diff(K, r')|$. We will prove that $r \sqsubseteq_K^{\trianglelefteq} r'$ iff $r \leqslant_K r'$ by induction on the size of $|Diff(K, r)|$.

<u>*Base Case*</u>: Assume that $|Diff(K, r)| = |Diff(K, r')| = 1$.

Then for some $p, q \in P$, $Diff(K, r) = \{p\}$ and $Diff(K, r') = \{q\}$. This again entails that $r \models \overline{p_K}$ and $r' \models \overline{q_K}$.

First assume that $r \sqsubseteq_K^{\trianglelefteq} r'$. Then by the definition of $\sqsubseteq_K^{\trianglelefteq}$ we derive that $p \trianglelefteq q$. Therefore, by the construction of $\trianglelefteq$, $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$. Moreover, by Lemma 3 it follows that r is the only $\leqslant_K$-minimal world in $[\overline{p_K}]$. Consequently $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$ entails that r is also $\leqslant_K$-minimal in $[\overline{p_K} \vee \overline{q_K}]$. Hence, from $r' \in [\overline{q_K}]$ we derive that $r \leqslant_K r'$ as desired. Thus we have shown that $r \sqsubseteq_K^{\trianglelefteq} r'$ entails $r \leqslant_K r'$.

For the converse, assume that $r \leqslant_K r'$. From Lemma 3 it follows that r is the only $\leqslant_K$-minimal world in $[\overline{p_K}]$, and r' is the only $\leqslant_K$-minimal world in $[\overline{q_K}]$. Hence, from

$r \preccurlyeq_K r'$ we derive that $p_K \notin K * (\overline{p_K} \vee \overline{q_K})$. Thus $p \trianglelefteq q$, and consequently $r \sqsubseteq^{\trianglelefteq}_K r'$ as desired.

Hence, the claim that $r \sqsubseteq^{\trianglelefteq}_K r'$ iff $r \preccurlyeq_K r'$ holds for $|\mathit{Diff}(K,r)| = |\mathit{Diff}(K,r')| = 1$.

<u>*Induction Hypothesis*</u>:

Assume that for $m \geqslant 1$, and any world $u, u' \in \mathcal{M}$, if $|\mathit{Diff}(K,r)| = |\mathit{Diff}(K,r')| = m$, then it holds that $u \sqsubseteq^{\trianglelefteq}_K u'$ iff $u \preccurlyeq_K u'$.

<u>*Induction step*</u>:

We will show that for any for any $r, r' \in \mathcal{M}$, if $|\mathit{Diff}(K,r)| = |\mathit{Diff}(K,r')| = m+1$, then it holds that $r \sqsubseteq^{\trianglelefteq}_K r'$ iff $r \preccurlyeq_K r'$.

Let $r, r' \in \mathcal{M}$ be such that $|\mathit{Diff}(K,r)| = |\mathit{Diff}(K,r')| = m+1$. We start by showing that $r \sqsubseteq^{\trianglelefteq}_K r'$ entails $r \preccurlyeq_K r'$. We distinguish between three cases.

First assume that $|\mathit{Diff}(K,r)_z| = |\mathit{Diff}(K,r')_z|$ for all $z \in P$. Let p and q be $\trianglelefteq$-*maximal* elements of $\mathit{Diff}(K,r)$ and $\mathit{Diff}(K,r')$ respectively. It is not hard to verify that $p \trianglelefteq q$ and $\mathit{Diff}(K,r) - \{p\} \trianglelefteq \mathit{Diff}(K,r') - \{q\}$.[13] Define u to be the world that differs from r only in p, and u' to be the world that differs from r' only in q. Then, $\mathit{Diff}(K,u) = \mathit{Diff}(K,r) - \{p\}$ and $\mathit{Diff}(K,u') = \mathit{Diff}(K,r') - \{q\}$. Hence, from $\mathit{Diff}(K,r) - \{p\} \trianglelefteq \mathit{Diff}(K,r') - \{q\}$, we derive that $u \sqsubseteq^{\trianglelefteq}_K u'$. Moreover, $|\mathit{Diff}(K,u)| = |\mathit{Diff}(K,u')| = m$ and therefore by the induction hypothesis, $u \preccurlyeq_K u'$. Call A, B the set of literals in K over the variables in $\mathit{Diff}(K,u)$ and $\mathit{Diff}(K,u')$ respectively. From Lemma 3 it follows that u is $\preccurlyeq_K$-minimal in $[\overline{A}]$, and u' is $\preccurlyeq_K$-minimal in $[\overline{B}]$. Hence, from $u \preccurlyeq_K u'$ it follows that $\neg(\overline{A}) \notin K * (\overline{A} \vee \overline{B})$. Consequently, since $p \trianglelefteq q$, from (D2) we derive that $\neg(\overline{A}\overline{p_K}) \notin K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})$. Hence all $\preccurlyeq_K$-minimal worlds in $[\overline{A}\overline{p_K}]$ belong to $[K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})]$. Observe that by the definition of A and Lemma 3, r is the only $\preccurlyeq_K$-minimal world in $[\overline{A}\overline{p}]$. Therefore, $r \in [K * (\overline{A}\overline{p} \vee \overline{B}\overline{q})]$. Consequently, since $r' \in [\overline{B}\overline{q}]$, we derive that $r \preccurlyeq_K r'$.

For the second case, assume that for some $z \in P$, $|\mathit{Diff}(K,r)_z| > |\mathit{Diff}(K,r')_z|$, and that $|\mathit{Diff}(K,r)_{z'}| = |\mathit{Diff}(K,r')_{z'}|$, for all $z' \triangleleft z$; moreover, assume that there exists

[13]Of course the converse also holds: $q \trianglelefteq p$ and $\mathit{Diff}(K,r') - \{q\} \trianglelefteq \mathit{Diff}(K,r) - \{p\}$.

$p \in \mathit{Diff}(K, r)$ such that $z \lhd p$.

Let q be a $\trianglelefteq$-maximal element in $\mathit{Diff}(K, r')$. Since $|\mathit{Diff}(K, r)| = |\mathit{Diff}(K, r')|$, by the assumptions of the case we derive that $z \lhd q$. Then it is not hard to verify that $\mathit{Diff}(K, r) - \{p\} \lhd \mathit{Diff}(K, r') - \{q\}$. Define u to be the world that differs from r only in p, and u to be the world that differs from r only in q. Then, $\mathit{Diff}(K, u) = \mathit{Diff}(K, r) - \{p\}$ and $\mathit{Diff}(K, u') = \mathit{Diff}(K, r') - \{q\}$. Hence, from $\mathit{Diff}(K, r) - \{p\} \lhd \mathit{Diff}(K, r') - \{q\}$, we derive that $u \sqsubset_K^{\trianglelefteq} u'$. Moreover, $|\mathit{Diff}(K, u)| = |\mathit{Diff}(K, u')| = m$, and therefore by the induction hypothesis, $u \prec_K u'$. Call A, B, the set of literals in K over the variables in $\mathit{Diff}(K, u)$ and $\mathit{Diff}(K, u')$ respectively. From Lemma 3 it follows that u is the only $\preccurlyeq_K$-minimal in $[\overline{A}]$, and u' is the only $\preccurlyeq_K$-minimal in $[\overline{B}]$. Hence, from $u \prec_K u'$ it follows that $\neg\overline{B} \in K * (\overline{A} \vee \overline{B})$. Moreover, since q is $\trianglelefteq$-maximal in $\mathit{Diff}(K, r')$, we derive that $z' \notin K * (\overline{z'} \vee \overline{q_K})$ for all $z' \in B$. Consequently from (D4), $\neg\overline{B}\overline{q_K} \in K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})$. Observe that by Lemma 3, r is the only $\preccurlyeq_K$-minimal worlds in $[\overline{A}\overline{p}]$ and r' is the only $\preccurlyeq_K$-minimal worlds in $[\overline{B}\overline{q}]$. Thus from $\neg\overline{B}\overline{q} \in K * (\overline{A}\overline{p} \vee \overline{B}\overline{q})$ we derive that $r \prec_K r'$.

For the third case, assume that for some $z \in P$, $|\mathit{Diff}(K, r)_z| > |\mathit{Diff}(K, r')_z|$, and that $|\mathit{Diff}(K, r)_{z'}| = |\mathit{Diff}(K, r')_{z'}|$, for all $z' \trianglelefteq z$; moreover assume that $p \trianglelefteq z$, for all $p \in \mathit{Diff}(K, r)$.

Clearly, z is maximal in $\mathit{Diff}(K, r)$. Define q to be the maximal element of $\mathit{Diff}(K, r')$. It is not hard to see that $z \lhd q$. Next we show that $\mathit{Diff}(K, r) - \{z\} \trianglelefteq \mathit{Diff}(K, r') - \{q\}$.

Observe that, if z is the only maximal element in $\mathit{Diff}(K, r)$, then $\mathit{Diff}(K, r) - \{z\} = \bigcup_{z' \lhd z} \mathit{Diff}(K, r)_{z'}$. Hence, since for all $z' \lhd z$ it holds that $|\mathit{Diff}(K, r)_{z'}| = |\mathit{Diff}(K, r')_{z'}|$, we derive that $\mathit{Diff}(K, r) - \{z\} \trianglelefteq \mathit{Diff}(K, r') - \{q\}$.[14] If on the other hand there is a $p \neq z$ such that $z \trianglelefteq p$ then, it is not hard to verify that from the assumptions of the case, it follows that $|(\mathit{Diff}(K, r) - \{z\})_{z'}| = |(\mathit{Diff}(K, r') - \{q\})_{z'}|$ for all $z' \lhd p$, and $|(\mathit{Diff}(K, r) - \{z\})_p| \geqslant |(\mathit{Diff}(K, r') - \{q\})_p|$. From this and $|\mathit{Diff}(K, r) - \{z\}| = |\mathit{Diff}(K, r') - \{q\}|$ we derive that once again $\mathit{Diff}(K, r) - \{z\} \trianglelefteq \mathit{Diff}(K, r') - \{q\}$.

Next we proceed in a fashion similar to the previous case. In particular, define u to

[14]Notice that for all $z' \lhd z$, $\mathit{Diff}(K, r)_{z'} = (\mathit{Diff}(K, r) - \{z\})_{z'}$ and $\mathit{Diff}(K, r')_{z'} = (\mathit{Diff}(K, r') - \{q\})_{z'}$.

be the world that differs from r only in z, and u to be the world that differs from r only in q. Then, $\mathit{Diff}(K,u) = \mathit{Diff}(K,r) - \{z\}$ and $\mathit{Diff}(K,u') = \mathit{Diff}(K,r') - \{q\}$. Hence, from $\mathit{Diff}(K,r) - \{z\} \trianglelefteq \mathit{Diff}(K,r') - \{q\}$, we derive that $u \sqsubseteq_K^{\trianglelefteq} u'$. Moreover, $|\mathit{Diff}(K,u)| = |\mathit{Diff}(K,u')| = m$, and therefore by the induction hypothesis, $u \preccurlyeq_K u'$. Call A, B, the set of literals in K over the variables in $\mathit{Diff}(K,u)$ and $\mathit{Diff}(K,u')$ respectively. From Lemma 3 it follows that u is $\preccurlyeq_K$-minimal in $[\overline{A}]$, and u' is $\preccurlyeq_K$-minimal in $[\overline{B}]$. Hence, from $u \preccurlyeq_K u'$ it follows that $\neg\overline{A} \notin K * (\overline{A} \vee \overline{B})$. Consequently, since $z \lhd q$, from (D3) we derive that $\neg(\overline{A}\overline{p_K}) \notin K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})$. Hence, $[K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})]$ contains a $\overline{A}\overline{p_K}$-world. Observe that by the definition of A and Lemma 3, r is a $\preccurlyeq_K$-minimal world in $[\overline{A}\overline{p_K}]$. Therefore, $r \in [K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})]$. Consequently, since $r' \in [\overline{B}\overline{q_K}]$, we derive that $r \preccurlyeq_K r'$.

We have thus shown that $r \sqsubseteq_K^{\trianglelefteq} r'$ entails $r \preccurlyeq_K r'$ (under the assumptions of the induction step).

For the converse, assume that $r \preccurlyeq_K r'$. Let p, q be $\trianglelefteq$-maximal elements in $\mathit{Diff}(K,r)$ and $\mathit{Diff}(K,r')$ respectively. Moreover, let A, B be the set of literals in K over $\mathit{Diff}(K,r) - \{p\}$ and over $\mathit{Diff}(K,r') - \{q\}$ respectively. From Lemma 3 it follows that r is $\preccurlyeq_K$-minimal in $[\overline{A}\overline{p_K}]$ and r' is $\preccurlyeq_K$-minimal in $[\overline{B}\overline{q_K}]$. Hence, since $r \preccurlyeq_K r'$ we derive that $\neg(\overline{A}\overline{p_K}) \notin K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})$. Moreover, since p is $\trianglelefteq$-maximal in $\mathit{Diff}(K,r)$, we have that for all $z \in \mathit{Diff}(K,r)$, $z_K \notin K * (\overline{z_K} \vee \overline{p_K})$. Consequently, from the counter-positive of (D4) we derive that $\neg\overline{A} \notin K * (\overline{A} \vee \overline{B})$.

Let u be the world that agrees with r over all variables except p, and let u' be the world that agrees with r' over all variables except q. Clearly, $u \in [\overline{A}]$ and $u' \in [\overline{B}]$. From Lemma 3 it follows that u is a $\preccurlyeq_K$-minimal world in $[\overline{A}]$. Hence from $\neg\overline{A} \notin K * (\overline{A} \vee \overline{B})$, we derive that u is also $\preccurlyeq_K$-minimal in $[\overline{A} \vee \overline{B}]$. Therefore, from $u' \in [\overline{B}]$ it follows that $u \preccurlyeq_K u'$. Consequently, by the induction hypothesis, $u \sqsubseteq_K^{\trianglelefteq} u'$. Hence, $\mathit{Diff}(K,u) \trianglelefteq \mathit{Diff}(K,u')$ or equivalently, $\mathit{Diff}(K,r) - \{p\} \trianglelefteq \mathit{Diff}(K,r') - \{q\}$. We proceed by distinguishing between three cases.

First, assume that $p \trianglelefteq q$ and $\mathit{Diff}(K,r') - \{q\} \trianglelefteq \mathit{Diff}(K,r) - \{p\}$. Since we have shown that $\mathit{Diff}(K,r) - \{p\} \trianglelefteq \mathit{Diff}(K,r') - \{q\}$, we derive that $|(\mathit{Diff}(K,r) - \{p\})_{z'}| =$

$|(Diff(K,r') - \{q\}_{z'})|$ for all $z' \in P$. Consequently, either $|Diff(K,r)_{z'}| = |Diff(K,r')_{z'}|$ for all $z' \in P$, or $|Diff(K,r)_p| > |Diff(K,r')_p|$ and $|Diff(K,r)_{z'}| = |Diff(K,r')_{z'}|$ for all $z' \lhd p$. In either case $Diff(K,r) \trianglelefteq Diff(K,r')$ and therefore $r \sqsubseteq_K^{\trianglelefteq} r'$.

For the second case, assume that $p \trianglelefteq q$ and $Diff(K,r) - \{p\} \lhd Diff(K,r') - \{q\}$. Then for some $z \in P$, $|(Diff(K,r) - \{p\})_z| > |(Diff(K,r') - \{q\})_z)|$ and $|(Diff(K,r) - \{p\})_{z'}| = |(Diff(K,r') - \{q\})_{z'})|$ for all $z' \lhd z$. If $z \lhd p$, or if p and q are equivalent wrt $\trianglelefteq$ (i.e., $p \trianglelefteq q$ and $q \trianglelefteq p$), we derive that $|Diff(K,r)_z| > |Diff(K,r')_z|$ and $|Diff(K,r)_{z'}| = |Diff(K,r')_{z'}|$ for all $z' \lhd z$; thus $r \sqsubseteq_K^{\trianglelefteq} r'$. If on the other hand $p \trianglelefteq z$ and $p \lhd q$, then $|Diff(K,r)_p| > |Diff(K,r')_p|$ and $|Diff(K,r)_{z'}| = |Diff(K,r')_{z'}|$ for all $z' \lhd p$. Hence, once again, $r \sqsubseteq_K^{\trianglelefteq} r'$.

For the third case, assume that $q \lhd p$. If $\neg\overline{B} \notin K * (\overline{A} \vee \overline{B})$, then from (D3) we derive that $\neg\overline{A}\overline{p}_K \in K * (\overline{A}\overline{p_K} \vee \overline{B}\overline{q_K})$. This however contradicts our initial assumption that $r \preccurlyeq_K r'$ (recall that by Lemma 3, r and r' are $\preccurlyeq_K$-minimal in $[\overline{A}\overline{p}]$ and $[\overline{B}\overline{q}]$ respectively). Hence we derive that $\neg\overline{B} \in K * (\overline{A} \vee \overline{B})$, and therefore $u \prec_K u'$. By the induction hypothesis we then derive that $u \sqsubset_K^{\trianglelefteq} u'$, and therefore $Diff(K,r) - \{p\} \lhd Diff(K,r') - \{q\}$. Hence, since q is $\trianglelefteq$-maximal in $Diff(K,r')$, it is not hard to verify that by the definition of $\trianglelefteq$, $Diff(K,r) \trianglelefteq Diff(K,r')$, and therefore $r \sqsubseteq_K^{\trianglelefteq} r'$. $\Box$

According to Theorem 3, if the revision function $*$ satisfies (D1) – (D4) at K, then there exists a preorder $\trianglelefteq$ over P, such that $\sqsubseteq_K^{\trianglelefteq}$ is identical to the preorder that $*$ assigns to K. Clearly, if $*$ also satisfies (D1) – (D4) at some other theory H, then Theorem 3 entails that the preorder that $*$ assigns to H can also be induced from some preorder $\trianglelefteq'$ over atoms. Notice however, that $\trianglelefteq$ and $\trianglelefteq'$ are not necessarily the same. To ensure this we need the axiom (D5) below:

(D5) If K, H are complete, $p \leqq_K q$, $x \in \{p, \overline{p}\}$, $y \in \{q, \overline{q}\}$ and $x, y \in H$, then $x \leqq_H y$.

Axiom (D5) says that if for a given theory K it is at least as easy to reverse p as it is to reverse q, then this relationship is preserved for any other theory H and any other two literals x, y that share the same atoms with p and q respectively; for example if $p \leqq_K q$ and $\neg p, q \in H$, then $\neg p \leqq_H q$.

It can be shown that the addition of (D5) to (D1) – (D4) suffices to characterise the AGM revision functions that assign PD preorders to every consistent complete theory, all of which are generated from the *same* preorder $\trianglelefteq$ over P.

Theorem 4. *If $*$ is a PD operator, then (D5) is satisfied for any two consistent complete theories K, H.*

Proof. Assume that $*$ is a PD operator, and let $\trianglelefteq$ be the preorder over atoms that induces $*$.

Next some notation that we shall use in this proof. For any literal x, we define the *variable of* x, denote x_v as follows: if $x \in P$ then $x_v = x$; if on the other hand $x = \neg y$ for some $y \in P$, then $x_v = y$.

Consider now any two consistent complete theories K, H and literals p, q, x, y such that $p, q \in K$, $x, y \in H$, $x \in \{p, \overline{p}\}$, $y \in \{q, \overline{q}\}$, and $p \notin K * (\overline{p} \vee \overline{q})$. We will show that $x \notin H * (\overline{x} \vee \overline{y})$.

From $p \notin K * (\overline{p} \vee \overline{q})$ it follows that there is a world $r \in [\overline{p}]$ such that r is $\sqsubseteq_K^{\trianglelefteq}$-minimal in $[\overline{p} \vee \overline{q}]$. Hence there is a world $w \in [K]$ such that $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w', r')$ for all $w' \in [K]$ and all $r' \in [\overline{q}]$. It is not hard to verify that r agrees with w over all variables except p_v; i.e., $\mathit{Diff}(w, r) = \{p_v\}$. Define r' to be the world that agrees with w over all variables except q_v; i.e., $\mathit{Diff}(w, r') = \{q_v\}$. Clearly, $r' \in [\overline{q}]$ and therefore $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w, r')$. Hence we derive that $p_v \trianglelefteq q_v$.

Since $x \in \{p, \overline{p}\}$ and $y \in \{q, \overline{q}\}$, it follows that $x_v \trianglelefteq y_v$. Let u be any world in $[H]$ and s the world that agrees with u over all variables except x_v. Then $s \in [\overline{x}]$ and $\mathit{Diff}(u, s) = \{x_v\}$. Moreover obverse that for any world $s' \in [\overline{x} \vee \overline{y}]$ and any $u' \in [H]$, $\mathit{Diff}(u', s')$ contains at least one of x_v or y_v. Hence from $x_v \trianglelefteq y_v$ we derive that $\mathit{Diff}(u, s) \trianglelefteq \mathit{Diff}(u', s')$ for all $u' \in [H]$ and $s' \in [\overline{x} \vee \overline{y}]$. This again entails that s is $\sqsubseteq_H^{\trianglelefteq}$-minimal in $[\overline{x} \vee \overline{y}]$ and therefore $x \notin H * (\overline{x} \vee \overline{y})$. $\square$

The following corollary follows immediately from Theorems 2, 3, and 4:

Corollary 1. *Let $*$ be an AGM revision function. Then $*$ satisfies (D1) – (D5) iff there* *exists a total preorder over atoms $\trianglelefteq$, such that $\sqsubseteq_K^{\trianglelefteq}$ is identical to the faithful preorder that $*$ assigns to K, for all consistent complete theories K.*

5.2. The General Case

We now turn to the general case of arbitrary consistent theories as belief sets. The characterisation of PD revision for this case requires some extra notation.

Consider any two possible worlds w, w' and a sentence $\varphi \in L$. By $\varphi(w, w')$ we denote the sentence produced from φ by replacing every variable in $Diff(w, w')$ with its negation.

For example, consider the worlds $w = abc$ and $w' = \overline{a}\overline{b}c$ (over the language built from the propositional variables a, b, and c). Clearly, $Diff(w, w') = \{a, b\}$. Hence for $\varphi = (\neg a \vee c) \wedge \neg b$, it follows that $\varphi(w, w') = (\neg\neg a \vee c) \wedge \neg\neg b$, which of course is equivalent to $(a \vee c) \wedge b$.

Intuitively, the above mapping translates any sentence φ into a sentence $\varphi(w, w')$ that "differs" from w' in exactly the same way that φ "differs" from w. This is made more precise by the following lemma:

Lemma 4. *Let w, w', r, r' be possible worlds such that $Diff(w, r) = Diff(w', r')$. Then for any contigent sentence $\varphi \in L$, $r \in [\varphi]$ iff $r' \in [\varphi(w, w')]$.*

Proof. Let φ be a consistent sentence of L. We prove the lemma by induction on the number of boolean operators in φ.

Firstly assume that φ has no boolean operators (*base case* of the induction). Hence φ is a propositional variable. If $\varphi \notin Diff(w, w')$, then $\varphi(w, w') = \varphi$, and therefore either both w, w' satisfy φ or they both falsify it. Consequently, since $Diff(w, r) = Diff(w', r')$, either both r, r' satisfy φ or they both falsify it. That is, $r \in [\varphi]$ iff $r' \in [\varphi(w, w')]$ as desired.

Assume that the lemma holds for any sentence φ with up to k boolean operators (*induction hypothesis*).

Consider now a sentence φ with $k + 1$ boolean operators (*induction step*). We distinguish between five cases depending on which of the boolean connective is the main connective of φ.

Suppose that the main connective of φ is negation, i.e. $\varphi = \neg\psi$ for some sentence $\psi \in L$. Clearly ψ has k connectives, it is contingent (since φ is contingent) and therefore by the induction hypothesis, $r \in [\psi]$ iff $r' \in [\psi(w, w')]$. Moreover, by definition, $(\neg\psi)(w, w') = \neg(\psi(w, w'))$. Consequently, $r \in [\neg\psi]$ iff $r' \in [(\neg\psi)(w, w')]$, or equivalently, $r \in [\varphi]$ iff $r' \in [\varphi(w, w')]$ as desired.

Next assume that the main connective is disjunction, i.e. $\varphi = \chi \vee \psi$ for some sentences $\chi, \psi \in L$. Clearly each of χ, ψ have at most k boolean connectives. Moreover, since φ is contingent, either both χ and ψ are contingent, or one of them is a contradiction and the other is a contingent sentence.

In the former case, by the induction hypothesis we derive that $r \in [\chi]$ iff $r' \in [\chi(w, w')]$ and $r \in [\psi]$ iff $r' \in [\psi(w, w')]$. Hence $r \in [\chi \vee \psi]$ iff $r' \in [\chi(w, w') \vee \psi(w, w')]$. Finally observe that by definition, $(\chi \vee \psi)(w, w') = \chi(w, w') \vee \psi(w, w')$. Therefore $r \in [\varphi]$ iff $r' \in [\varphi(w, w')]$ as desired.

In the latter case, we can assume without loss of generality that χ is a contradiction and ψ is contingent. By the induction hypothesis we get that $r \in [\psi]$ iff $r' \in [\psi(w, w')]$. Moreover it is not hard to verify that since χ is inconsistent replacing any variable in χ with its negation preserves inconsistency. Therefore $\chi(w, w')$ is also inconsistent. Consequently it follows that $r' \in [\chi(w, w') \vee \psi(w, w')]$. Hence since $(\chi \vee \psi)(w, w') = \chi(w, w') \vee \psi(w, w')$, we derive that $r \in [\varphi]$ iff $r' \in [\varphi(w, w')]$ as desired.

The remaining three cases where the main connective of φ is conjunction, implication, and equivalence, are totally analogous to the previous case. $\square$

Lemma 4 entails immediately the following corollary which confirms more explicitly our statement that w differs from φ in exactly the same way that w' differs from $\varphi(w, w')$; or equivalently, that $\varphi(w, w')$ is, loosely speaking, what φ looks like when our point of view changes from w to w':

Corollary 2. *Let w, w' be any two possible worlds and $\varphi \in L$ a contingent sentence. For each $r \in [\varphi]$ there is a $r' \in [\varphi(w, w')]$ such that $\mathit{Diff}(w, r) = \mathit{Diff}(w', r')$; and conversely, for each $r' \in [\varphi(w, w')]$ there is a $r \in [\varphi]$ such that $\mathit{Diff}(w, r) = \mathit{Diff}(w', r')$.*

The following property is the only one we need to add to (D1) – (D5) in order to characterise PD revision for the general case:

$$\text{(D6)}\quad K * \varphi = Cn(\vee \left\{ \begin{array}{ll} \psi \in L : & \text{for some } w \in [K],\ Cn(\psi) = w * \varphi, \text{and for all} \\ & w' \in [K],\ \neg\psi(w, w') \notin w' * (\varphi \vee \psi(w, w')) \end{array} \right\})$$

Conditions (D6) defines the revision by φ of any consistent theory K, in terms of the revision by φ of the possible worlds w of K (i.e. $w \in [K]$). Loosely speaking, (D6) says that $K * \varphi$ is the disjunction of all $w * \varphi$ with $w \in [K]$, for which there is no $w' \in [K]$ such that φ is strictly more plausible than $w * \varphi$ when the latter is seen from w''s point of view.

Theorem 5. *Let $*$ be a PD operator and K a consistent theory of L. Then for all consistent sentences $\varphi \in L$, (D6) is satisfied.*

Proof. Since $*$ is a PD operator, there exists a preorder $\trianglelefteq$ over the atoms of L that determines the faithful preorder assigned to K.

Let $\varphi \in L$ be any consistent sentence. We prove the theorem by showing that $[K * \varphi]$ = $[\vee\{\psi \in L$: for some $w \in [K]$, $Cn(\psi) = w * \varphi$, and for all $w' \in [K]$, $\neg\psi(w, w') \notin w' * (\varphi \vee \psi(w, w'))\}]$.

LHS $\subseteq$ RHS

Consider any $r \in [K * \varphi]$. Then $r \in [\varphi]$ and moreover, there exists a $z \in [K]$ such that $\mathit{Diff}(z, r) \trianglelefteq \mathit{Diff}(z', r')$, for all $z' \in [K]$ and $r' \in [\varphi]$; i.e. $\mathit{Diff}(z, r)$ is $\trianglelefteq$-minimal among all $\mathit{Diff}(z', r')$ for which $z' \in [K]$ and $r' \in [\varphi]$. Clearly then $r \in [z * \varphi]$.

Let ψ be any sentence such that $Cn(\psi) = z * \varphi$[15] and assume towards contradiction that $r \notin [\vee\{\psi \in L$: for some $w \in [K]$, $Cn(\psi) = w * \varphi$, and for all $w' \in [K]$, $\neg\psi(w, w') \notin$

[15] Since L is assumed to be a propositional language built from *finitely many* propositional variables, such a sentence ψ always exists

$w' * (\varphi \vee \psi(w, w'))\}]$. Then, since $r \in [z * \varphi]$ and $z \in [K]$, we derive that there exists a $z' \in [K]$ such that $\neg\psi(z, z') \in z' * (\varphi \vee \psi(z, z'))\}]$. This in turn entails that there is a world $u' \in [\varphi]$ such that $\mathit{Diff}(z', u') \lhd \mathit{Diff}(z'', u'')$ for all $z'' \in [K]$ and $u'' \in [\psi(z, z')]$. In particular, $\mathit{Diff}(z', u') \lhd \mathit{Diff}(z', u'')$, for all $u'' \in [\psi(z, z')]$. This however, in view of Corollary 2, leads us to a contradiction. Indeed, from Corollary 2 we derive that there is a world $r'' \in [\psi(z, z')]$ such that $\mathit{Diff}(z, r) = \mathit{Diff}(z', r'')$. Combining the above it follows that $\mathit{Diff}(z', u') \lhd \mathit{Diff}(z, r)$. Given that $z' \in [K]$ and $u' \in [\varphi]$, this contradicts our initial assumption about the $\trianglelefteq$-minimality of $\mathit{Diff}(z, r)$.

RHS $\subseteq$ LHS
Let r be any world in $[\vee\{\psi \in L$: for some $w \in [K]$, $Cn(\psi) = w * \varphi$, and for all $w' \in [K]$, $\neg\psi(w, w') \notin w' * (\varphi \vee \psi(w, w'))\}]$.

Then there exists $\psi \in L$ such that for some $w \in [K]$, $Cn(\psi) = w * \varphi$, $r \in [w * \varphi]$ and for all $w' \in [K]$, $\neg\psi(w, w') \notin w' * (\varphi \vee \psi(w, w'))$. From $r \in [w * \varphi]$ we derive that, firstly, $r \in [\varphi]$, and moreover $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w, r'')$ $r'' \in [\varphi]$. Likewise, from $\neg\psi(w, w') \notin w' * (\varphi \vee \psi(w, w'))$ we derive that there is a $u \in [\psi(w, w')]$ such that $\mathit{Diff}(w', u) \trianglelefteq \mathit{Diff}(w'', u'')$ for all $w'' \in [K]$ and $u'' \in [\varphi]$. Finally, from Collorary 2 we then derive that there is a $r' \in [\varphi]$ such that $\mathit{Diff}(w, r') = \mathit{Diff}(w', u)$. Therefore, $\mathit{Diff}(w, r') \trianglelefteq \mathit{Diff}(w'', u'')$ for all $w'' \in [K]$ and $u'' \in [\varphi]$. Moreover, since $r' \in [\varphi]$ it follows that $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w, r')$. Hence $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w'', u'')$ for all $w'' \in [K]$ and $u'' \in [\varphi]$, which again entails that $r \in [K * \varphi]$. □

Theorem 6. *An AGM revision function $*$ is a PD operator iff it satisfies (D1) – (D6).*

Proof. The fact that all PD operators satisfy (D1) – (D6) follows from previous results.

For the converse, assume that (D1) – (D6) are satisfied. In this case we have already shown that the restriction of $*$ to complete theories is a PD operator. That is, there exists a preorder $\trianglelefteq$ over the atoms of L, such that the faithful preorders assigned by $*$ to any consistent complete theory is the one induced from $\trianglelefteq$. To complete the proof we need to show that this is also the case for any incomplete theory K. In particular, it suffices to show that for any sentence φ, the revision of K by φ as specified by (D6) is identical to revising K by φ in accordance to the faithful preorder induced by $\trianglelefteq$ at K. This however follows from Theorem 5. $\square$

6. Properties of PD Operators

In this section we look at some of the properties of PD operators, starting with their relationship with Parikh's notion of relevance-sensitive belief revision.

In [15], Parikh introduced a new postulate for belief revision, called postulate (P), to capture the intuition that when revising her belief, a rational agent changes only the part of her belief set that is relevant to the new information.

(P) If $K = Cn(\chi, \psi)$ where χ, ψ are sentences of disjoint sublanguages L_1, L_2 respectively, and $\phi \in L_1$, then $K * \phi = (Cn_{\mathcal{L}_1}(\chi) \circ \phi) + \psi$, where $\circ$ is a revision operator of the sublanguage L_1.

The intuition behind (P) should be obvious: when a belief set can be split into two (syntactically) disjoint compartments, and the new information ϕ can be expressed in terms of the language of the first compartment alone, then it is only the first compartment that is revised by ϕ (considered the one relevant to ϕ); the beliefs in the second compartment remain unaffected.

Postulate (P) was further analysed in [22] and two different interpretations of it were identified, called the *weak* and the *strong* version of (P). The weak version of postulate (P), which we denote (wP), is much more general and intuitive, and it is this version we shall use herein.

Before presenting (wP) we need some more notation: for any sentence x, L_x denotes the (unique) smallest language in which x can be expressed. Moreover, $\overline{L_x}$ denotes the complement language, that is the language built from the propositional variables that do not appear in L_x. With this additional notation we can now present (wP):

(wP) If $K = Cn(\{x, y\})$, $L_x \cap L_y = \varnothing$, and $\varphi \in L_x$, then $(K * \varphi) \cap \overline{L_x} = K \cap \overline{L_x}$.

Postulate (wP) is weaker than (P) since it only requires that the non-relevant part of K remains unchanged; it makes no commitment on how the relevant part of K is affected.

In [22], (wP) was characterised semantically in terms of constraints over faithful preorders.

First however, we recall some additional notation and definitions from [22].

For any nonempty set of propositional variables $S \subseteq P$, by L^S we shall denote the propositional language built from the variables in S. Moreover, we extend in the following the definition of difference, *Diff*, between worlds, to include the difference between an arbitrary consistent *theory* K and a world r.

In particular, consider now a consistent theory K, and let $Q = \{Q_1, \ldots, Q_n\}$ be a partition of P; i.e., $\bigcup Q = P$, $Q_i \neq \emptyset$, and $Q_i \cap Q_j = \emptyset$, for all $1 \leqslant i \neq j \leqslant n$. We say that $Q = \{Q_1, \ldots, Q_n\}$ is a K-*splitting* iff there exist sentences $\phi_1 \in L^{Q_1}, \ldots,$ $\phi_n \in L^{Q_n}$, such that $K = Cn(\{\phi_1, \ldots, \phi_n\})$. Parikh has shown in [15] that for every theory K there is a unique *finest* K-splitting, i.e., one which refines every other K-splitting.[16]

In [22], the difference between an arbitrary consistent theory K and a world r was defined using the finest splitting of K, call it F, as follows: $Diff(K, r) = \bigcup\{F_i \in F :$ for some $\phi \in L^{F_i}$, $K \models \phi$ and $r \models \neg\phi\}$ (see [22] for a detailed discussion on this definition).

With the extended definition of *Diff*, it was shown in [22] that (wP) can be semantically characterised by the following two constraints:

(Q1) If $Diff(K, r) \subset Diff(K, r')$ and $Diff(r, r') \cap Diff(K, r) = \emptyset$, then $r \prec r'$.

(Q2) If $Diff(K, r) = Diff(K, r')$ and $Diff(r, r') \cap Diff(K, r) = \emptyset$, then $r \approx r'$.

Theorem 7. *[22]. Let $*$ be a revision function satisfying (K*1) - (K*8), K a consistent theory, and $\preccurlyeq_K$ a preorder faithful to K, that corresponds to $*$ at K by means of ($\preccurlyeq *$).*

[16] A partition Q' refines another partition Q, iff for every $Q'_i \in Q'$ there is $Q_j \in Q$, such that $Q'_i \subseteq Q_j$.

Then $*$ satisfies (wP) at K iff $\preccurlyeq_K$ satisfies (Q1) - (Q2).

It was furthermore shown in [22] that (wP) is consistent with the AGM postulates (K*1) - (K*8).

Next we show that all PD operators satisfy (wP). To this aim we recall the following lemma from [15]:

Lemma A [15]. *Let K be a theory and $\{Q_1, \ldots, Q_n\}$ a partition of P. If $\{Q_1, \ldots, Q_n\}$ is a K–splitting, then for any $r_1, \ldots, r_n \in [K]$, $Mix(r_1, \ldots, r_n; Q_1, \ldots, Q_n)$ belongs to $[K]$. Conversely, if $Mix(r_1, \ldots, r_n; Q_1, \ldots, Q_n)$ belongs to $[K]$ for all $r_1, \ldots, r_n \in [K]$, then $\{Q_1, \ldots, Q_n\}$ is a K–splitting.*

In the lemma above, $Mix(r_1, \ldots, r_n; Q_1, \ldots, Q_n)$ denotes the unique world r that agrees with r_1 on the variables in Q_1, with r_2 on the variables in Q_2, …, and with r_n on the variables in Q_n.

We can now prove the theorem alluded earlier.

Theorem 8. *Let $*$ be a PD operator and K a consistent theory of L. Then for all consistent $\phi \in L$, (wP) is satisfied.*

Proof Let $\trianglelefteq$ be the total preorder over atoms that induces $*$. By definition, the faithful preorder that $*$ assigns to K is $\sqsubseteq_K^{\trianglelefteq}$.

In view of Theorem 7, to prove (wP) it suffices to show that $\sqsubseteq_K^{\trianglelefteq}$ satisfies (Q1) – (Q2). The proof of (Q1) follows the same line of reasoning used in the proof of Theorem 7 in [22]; the proof of (Q2) is somewhat different.

Starting with condition (Q1), let r, r' be any two possible worlds such that $\mathit{Diff}(K, r) \subset \mathit{Diff}(K, r')$ and $\mathit{Diff}(r, r') \cap \mathit{Diff}(K, r) = \varnothing$. Then clearly, $P - \mathit{Diff}(K, r) \neq \varnothing$. Let u be a world in $[K]$ that agrees with r on all variables in $P - \mathit{Diff}(K, r)$.[17] Moreover, let

[17] To see that such a world indeed exists, consider the sentence ψ defined as the conjunction of all literals in r that are built from variables in $P - \mathit{Diff}(K, r)$. Clearly then, $r \models \psi$. Moreover, $\neg\psi \notin K$, for otherwise $\mathit{Diff}(K, r)$ would include variables from $P - \mathit{Diff}(K, r)$, which is of course a contradiction. Hence there is a

z be a K-world that differs in the least number of variables from r when restricted to $\mathit{Diff}(K,r)$; i.e., $|\mathit{Diff}(z,r) \cap \mathit{Diff}(K,r)| \leqslant |\mathit{Diff}(z',r) \cap \mathit{Diff}(K,r)|$ for all $z' \in [K]$. Define w to be the world that agrees with z on the variables in $\mathit{Diff}(K,r)$ and agrees with u on the remaining variables. Clearly, $\mathit{Diff}(w,r) \subseteq \mathit{Diff}(K,r)$. Moreover, by the definition of Diff, $\{\mathit{Diff}(K,r), P - \mathit{Diff}(K,r)\}$ is a K-splitting, and consequently from Lemma A and the fact that $z, u \in [K]$, we derive that $w \in [K]$.

Consider now any world $w' \in [K]$. Since $\mathit{Diff}(K,r) \subset \mathit{Diff}(K,r')$, there is at least one sentence μ, built entirely from variables in $P - \mathit{Diff}(K,r)$, such that $K \models \mu$ and $r' \models \neg\mu$. Hence from $w' \in [K]$, we derive that $\mathit{Diff}(w',r') \cap (P - \mathit{Diff}(K,r)) \neq \varnothing$ and consequently, $|\mathit{Diff}(w',r') \cap (P - \mathit{Diff}(K,r))| > 0$. Moreover, since r and r' agree on the variables in $\mathit{Diff}(K,r)$ it follows that $|\mathit{Diff}(w',r') \cap \mathit{Diff}(K,r))| = |\mathit{Diff}(w',r) \cap \mathit{Diff}(K,r))|$. Consequently, $|\mathit{Diff}(w',r')| = |\mathit{Diff}(w',r') \cap \mathit{Diff}(K,r))| + |\mathit{Diff}(w',r') \cap (P - \mathit{Diff}(K,r))| > |\mathit{Diff}(w',r') \cap \mathit{Diff}(K,r))| = |\mathit{Diff}(w',r) \cap \mathit{Diff}(K,r))| \geqslant |\mathit{Diff}(w,r) \cap \mathit{Diff}(K,r))| = |\mathit{Diff}(w,r)|$. Hence we have shown that $|\mathit{Diff}(w,r)| < |\mathit{Diff}(w',r')|$ for all $w' \in [K]$. This again entails that $r \sqsubset^{\trianglelefteq}_{K} r'$ as desired.

For (Q2), let K be a consistent theory and let r, r' be any two possible worlds such that $\mathit{Diff}(K,r) = \mathit{Diff}(K,r')$ and $\mathit{Diff}(r,r') \cap \mathit{Diff}(K,r) = \varnothing$. If $\mathit{Diff}(K,r) = P$ then $r = r'$ and (Q2) trivially holds. Moreover, if $\mathit{Diff}(K,r) = \varnothing$, then $r, r' \in [K]$ and therefore (Q2) follows from the fact that $\sqsubseteq^{\trianglelefteq}_{K}$ is faithful to K. Assume therefore that $\varnothing \neq \mathit{Diff}(K,r) \subset P$.

Let Q be the set, $Q = \{\mathit{Diff}(u,r) : u \in [K]\}$ and let $\mathit{Diff}(w,r)$ be a minimal element of Q with respect of $\trianglelefteq$; i.e., $w \in [K]$ and for all $u \in [K]$, $\mathit{Diff}(w,r) \trianglelefteq \mathit{Diff}(u,r)$.

Next we show that $\mathit{Diff}(w,r) \subseteq \mathit{Diff}(K,r)$. Assume on the contrary that $\mathit{Diff}(w,r) \cap (P - \mathit{Diff}(K,r)) \neq \varnothing$. Since r does not differ from K in any variables in $P - \mathit{Diff}(K,r)$, we derive that there is a $v \in [K]$ that agrees with r on all variables in $P - \mathit{Diff}(K,r)$ (see Footnote 17). Define z to be the world that agrees with w on the variables in $\mathit{Diff}(K,r)$ and agrees with v on the remaining variables. Then $\mathit{Diff}(z,r) \subset \mathit{Diff}(w,r)$

$u \in [K]$ such that $u \models \psi$. By the construction of ψ it follows that u agrees with r on all variables outside $\mathit{Diff}(K,r)$.

and moreover, since $\{\mathit{Diff}(K,r), P - \mathit{Diff}(K,r)\}$ is a K-splitting, by Lemma A, $z \in [K]$. This however contradicts our assumption that $\mathit{Diff}(w,r)$ is $\trianglelefteq$-minimal in Q. Hence we have shown that $\mathit{Diff}(w,r) \subseteq \mathit{Diff}(K,r)$; i.e., w agrees with r over all variables in $P - \mathit{Diff}(K,r)$.

Now pick a world $w' \in [K]$ such that $\mathit{Diff}(w',r')$ is $\trianglelefteq$-minimal in the set $Q' = \{\mathit{Diff}(u,r') : u \in [K]\}$. By a similar argument as the one above, we derive that $\mathit{Diff}(w',r') \subseteq \mathit{Diff}(K,r)$.

Next we show that $\mathit{Diff}(w,r) \trianglelefteq \mathit{Diff}(w',r')$. Assume towards contradiction that $\mathit{Diff}(w',r') \lhd \mathit{Diff}(w,r)$. Define z to be the world that agrees with w' over the variables in $\mathit{Diff}(K,r)$, and it agrees with w over all remaining variables. By Lemma A, $z \in [K]$. Moreover by construction, $\mathit{Diff}(z,r) \subseteq \mathit{Diff}(K,r)$. Hence, since r and r' agree over the variables in $\mathit{Diff}(K,r)$, and so do z and w', we derive that $\mathit{Diff}(z,r) = \mathit{Diff}(w',r')$. From $\mathit{Diff}(w',r') \lhd \mathit{Diff}(w,r)$ we then derive that $\mathit{Diff}(z,r) \lhd \mathit{Diff}(w,r)$. This of course contradicts our initial assumption that $\mathit{Diff}(w,r)$ is $\trianglelefteq$-minimal in $\{\mathit{Diff}(u,r) : u \in [K]\}$.

Thus we have shown that $\mathit{Diff}(w,r) \trianglelefteq \mathit{Diff}(w',r')$. Since $\mathit{Diff}(w',r')$ is $\trianglelefteq$-minimal in $\{\mathit{Diff}(u,r') : u \in [K]\}$ we derive that $\mathit{Diff}(w,r) \trianglelefteq \mathit{Diff}(u,r')$, for all $u \in [K]$. Consequently, $r \sqsubseteq_K^{\trianglelefteq} r'$.

By a totally symmetric argument we also derive that $r' \sqsubseteq_K^{\trianglelefteq} r$, thus proving (Q2). □

The next two properties satisfied by PD operators make connections between the revision policies related to *different* theories. They arise from the fact that all faithful preorders (over different theories) related to a PD operator $*$, are generated by a *common* preorder over atoms $\trianglelefteq$:

(N1) If $\neg\varphi \notin K * (\varphi \vee \psi)$ and $\neg\varphi \notin H * (\varphi \vee \psi)$, then $\neg\varphi \notin (K \cap H) * (\varphi \vee \psi)$.

(N2) If $\neg\varphi \in K * (\varphi \vee \psi)$ and $\neg\varphi \in H * (\varphi \vee \psi)$, then $\neg\varphi \in (K \cap H) * (\varphi \vee \psi)$.

Both conditions are intuitive and easy to understand: (N1) says that if bringing about φ is at least as easy as bringing about ψ, regardless of whether one starts at K or at H, then this is also true when one starts at the belief set containing the beliefs that are common in K and H.

Condition (N2) is drawn from the same intuition: if it is (strictly) easier to bring about ψ than it is to if bringing about φ, regardless of whether the initial belief set is K or H, then this is also true for the belief set containing the common beliefs of K and H.

Theorem 9. *Let $*$ be a PD operator and K, H consistent theories of L. Then for all consistent $\varphi, \psi \in L$, conditions (N1) – (N2) are satisfied.*

Proof. Let $\trianglelefteq$ be the total preorder over atoms that generates the PD operator $*$. Then, by definition, the faithful preorders that $*$ assigns to K and H are $\sqsubseteq_K^{\trianglelefteq}$ and $\sqsubseteq_H^{\trianglelefteq}$ respectively.

For condition (N1), let φ, ψ be consistent sentences in L such that $\neg\varphi \notin K * (\varphi \vee \psi)$ and $\neg\varphi \notin H * (\varphi \vee \psi)$. Then there exist worlds $r, u \in [\varphi]$ such that $r \sqsubseteq_K^{\trianglelefteq} v$ and $u \sqsubseteq_H^{\trianglelefteq} v$ for all $v \in [\psi]$. Consequently, there exist worlds $w \in [K]$ and $z \in [H]$ such that $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w', v)$, and $\mathit{Diff}(z, u) \trianglelefteq \mathit{Diff}(z', v)$ for all $w' \in [K]$, $z' \in [H]$, and $v \in [\psi]$. Since $\trianglelefteq$ is total, $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(z, u)$ or $\mathit{Diff}(z, u) \trianglelefteq \mathit{Diff}(w, r)$. Without loss of generality we assume the former. Then, $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w', v)$, for all $w' \in [K] \cup [H]$ and all $v \in [\psi]$. Consequently $r \sqsubseteq_{K \cap H}^{\trianglelefteq} v$, for all $v \in [\psi]$. Since by definition, $\sqsubseteq_{K \cap H}^{\trianglelefteq}$ is the faithful preorder that $*$ assigns to $K \cap H$, we derive that $\neg\varphi \notin (K \cap H) * (\varphi \vee \psi)$ as desired.

The proof of (N2) is very similar. In particular, let φ, ψ be consistent sentences in L such that $\neg\varphi \in K * (\varphi \vee \psi)$ and $\neg\varphi \in H * (\varphi \vee \psi)$. Then there exist worlds $r, u \in [\psi]$ such that $r \sqsubset_K^{\trianglelefteq} v$ and $u \sqsubset_H^{\trianglelefteq} v$ for all $v \in [\varphi]$. Consequently, there exist worlds $w \in [K]$ and $z \in [H]$ such that $\mathit{Diff}(w, r) \triangleleft \mathit{Diff}(w', v)$, and $\mathit{Diff}(z, u) \triangleleft \mathit{Diff}(z', v)$ for all $w' \in [K]$, $z' \in [H]$, and $v \in [\varphi]$. Since $\trianglelefteq$ is total, $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(z, u)$ or $\mathit{Diff}(z, u) \trianglelefteq \mathit{Diff}(w, r)$. Without loss of generality we assume the former. Then, $\mathit{Diff}(w, r) \trianglelefteq \mathit{Diff}(w', v)$, for all $w' \in [K] \cup [H]$ and all $v \in [\varphi]$. Consequently $r \sqsubset_{K \cap H}^{\trianglelefteq} v$, for all $v \in [\varphi]$. Since by definition, $\sqsubseteq_{K \cap H}^{\trianglelefteq}$ is the faithful preorder that $*$ assigns to $K \cap H$, we derive that $\neg\varphi \in (K \cap H) * (\varphi \vee \psi)$ as desired. $\Box$

We conclude this section with an axiomatic characterisation of Dalal's operator which as already stated, has hitherto been perhaps the most popular concrete revision operator in the literature.

Clearly, Dalal's operator is a PD operator. Hence to characterise Dalal's operator it suffices to add to the conditions (D1) – (D6) the specific feature that separates it from the remaining PD operators; namely, Dalal's operator assigns the *same* epistemic value to all atoms of the language: $a \trianglelefteq^D b$ for all $a, b \in P$.

Axiomatically, this can be expressed as follows:

(DL) For any propositional variables $a, b \in K$, $a \notin K * (\neg a \vee \neg b)$.

Theorem 10. *Let $*$ be an AGM revision function. Then $*$ is Dalal's operator iff $*$ satisfies (D1) – (D6) and (DL).*

Proof.

$(\Rightarrow)$

Assume that $*$ is Dalal's operator. Then $*$ is a PD operator and hence by Theorem 6, it satisfies (D1) – (D6).

For (DL), assume that K is a consistent theory and a, b are propositional variables that belong to K. Let w be any world in $[K]$. Clearly, $w \models a, b$. Define r to the world that agrees with w in all atoms except a. Then, $r \models \neg a$ and $|Diff(w, r)| = 1$.

Consider now any world $r' \in [\neg a \vee \neg b]$ and any $w' \in [K]$. Since $r' \notin [K]$, it follows that $|Diff(w', r')| \geqslant 1$. Consequently, $|Diff(w, r)| \leqslant |Diff(w', r')|$ for all $w' \in [K]$ and $r' \in [\neg a \vee \neg b]$. This makes r minimal in $[\neg a \vee \neg b]$ with respect to Dalal's preorder $\sqsubseteq_K$. Hence $r \in [K * (\neg a \vee \neg b)]$, which again entails that $a \notin K * (\neg a \vee \neg b)$ as desired.

$(\Leftarrow)$

Assume that $*$ satisfies (D1) – (D6) and (DL). Then by Theorem 6, $*$ is a PD operator, and therefore there exists a total preorder $\trianglelefteq$ over atoms that induces a family of preorders $\{\sqsubseteq_K^{\trianglelefteq}\}_{K \in \mathcal{K}}$, which in turn induces $*$. Given the construction of $\{\sqsubseteq_K^{\trianglelefteq}\}_{K \in \mathcal{K}}$ from $\trianglelefteq$, in order to complete the proof it suffices to show that all atoms have the same epistemic value with respect to $\trianglelefteq$.

Assume towards contradiction that for some $a, b \in P$, $b \lhd a$. Let w be any possible world in $[a \wedge b]$. Define r to be the world that differs from w only in b. Clearly then by construction, $\mathit{Diff}(w, r) \lhd \mathit{Diff}(w, r')$ for all $r' \in [\neg b] - \{r\}$ and $\mathit{Diff}(w, r) \lhd \mathit{Diff}(w, r'')$ for all $r'' \in [\neg a]$. Hence r is the only minimal world in $\{\mathit{Diff}(w, z) : z \in [\neg a] \cup [\neg b]\}$. Consequently, $w * (\neg a \vee \neg b) = r$. This however contradicts (DL), since by construction $a \in r$. Hence $a \trianglelefteq b$ for all $a, b \in P$, and therefore $*$ is Dalal's operator as desired. □

7. Implementation Considerations and Previous Work

In this section we will briefly consider what would be involved in an implementation of AGM belief revision. We also review previous work on this topic, and we sketch the benefits of using PD operators for such an implementation.[18]

An AGM belief revision solver would presumably answer queries of the form "does ψ hold after the revision by φ?". These queries will be assessed against a background *knowledge base* B[19] and a revision policy associated with B. Revision policies can be modelled in different ways, however they are typically encoded as preorders $\preccurlyeq$ either over possible worlds (faithful preorders), or over sentences (epistemic entrenchments), or sets of sentences (remainders). The problem is that the size of these preorders is, in general, *exponential* to the number of atoms in the object language. This high *representational cost* is one of the main obstacles in the development of real-world belief revision applications.

PD operators provide a very efficient solution to this problem: a single preorder $\trianglelefteq$ over the atoms of the object language L (hence *linear* in the number of atoms) suffices to determine the revision policy of *every* theory (or knowledge base) of L.

Observe that an added benefit of having a single preorder $\trianglelefteq$ generating the revision policy for all theories is that we thus also solve the problem of *iterated revision* with

[18] We note that implementing a belief revision solver based on PD operators is outside the scope of this article. See however [23] [24] for two such recent implementations.

[19] A knowledge base B is a finite set of sentences representing the initial belief set K; i.e., $K = Cn(B)$.

no extra representational cost: the revision policy at $K * \varphi$ is fully determined by $\trianglelefteq$, in the same way it is determined for any other theory.

A similar approach has been used in [25], where the revision policy at any theory, is determined by a single hierarchical order over all *formulas* of the language. This work was further extended in [26], [27], and [28].

Previous work on computational approaches to AGM revision, called *belief base revision schemes* in [29], are primarily syntax-based. We briefly review two of the most influential such approaches and compare them to PD operators.[20]

The first one is based on the notion of *ensconcement* introduced in [31].

Formally, an ensconcement $\leqslant$ related to a belief base B is defined as a preorder over the elements of B that satisfies the following constraints for all $\varphi, \psi \in B$:[21]

(i) If $\not\models \varphi$, then $\{\psi \in B : \varphi < \psi\} \not\models \varphi$.

(ii) If $\not\models \varphi$ and $\models \psi$, then $\varphi \leqslant \psi$ and $\psi \nleqslant \varphi$.

(ii) If $\models \varphi$ and $\models \psi$, then $\varphi \leqslant \psi$.

Intuitively, an ensconcement can be thought of as a succinct representation of an epistemic entrenchment. Indeed, it was shown in [31] that any ensconcement $\leqslant$ over B can be extended to an epistemic entrenchment related to $Cn(B)$. Moreover, it is possible to answer queries about the revision of $Cn(B)$, working directly with the ensconcement $\leqslant$, rather than the induced epistemic entrenchment. This addresses the problem of the representational cost, since the size of an ensconcement is linear to the size of the knowledge base B.

A second influential belief base revision scheme, called *linear belief base revision*, was introduced in [32]. This approach partitions a knowledge base B into priority classes $B_1, \ldots, B_n$. To revise B by a sentence φ, one removes an entire priority class B_i

[20] See [30] for more approaches that are likewise based on prioritised (belief) bases.

[21] The symbol $\leqslant$ that we use herein for ensconcement will be used in the next section to compare numbers; any ambiguity with this slight abuse of notation is resolved by the context.

if (one or more of) its sentences are responsible for a contradiction with φ, and none of the lower priority classes can be blamed for the contradiction. It was shown in [32] that this procedure induces revision functions that satisfy all the AGM postulates for revision. Moreover, this approach also deals with the problem of the representational cost since any partition of B is linear to the size of B.

Comparing ensconcement-based revision and linear belief base revision with PD operators, one can immediately identify two advantages of the latter. Firstly the size of a knowledge base is typically much larger than the number of atoms, and therefore PD revision has (in principle) a lower representational cost. Secondly, and more importantly, PD revision has an embedded solution to the iterated revision problem (at no extra representational cost). This is missing from both ensconcement-based revision and linear belief base revision: in both cases, new preorders need to be provided explicitly after each revision step (clearly a prohibitive requirement for real-world applications).

On the other hand, the formal results in [32], [31] seem to suggest that ensconcement-based revision and linear belief base revision have a greater *range of applicability* than PD operators. In particular, it has been shown that both these approaches can encode *any* AGM revision function. In contrast, the class of PD operators is a proper subclass of AGM revision functions (namely the subclass satisfying (D1) – (D6)). However a careful reading of the results in [32], [31] reveals a somewhat different picture.

It is true that any AGM revision function can be generated from prioritised knowledge bases with the method described by Nebel, but only if the belief base B is allowed to *vary* according to the desired revision policy. More precisely, given a knowledge base B and an AGM revision function $*$, it could well be the case that no prioritisation of B produces the same results as $*$ at $Cn(B)$. All that the results in [32] tell us is that, in that case, there exists some *other* belief base B' that is logically equivalent to B, for which such a prioritisation can be found.

Yet, we argue, that a knowledge base B ought to be independent from the revision policy employed. Adding to B (logically) redundant sentences, just to address the technical requirements of a certain revision representation method, is not in our view an elegant way to increase the range of applicability.

The results in [31] also require a varying knowledge base, and therefore the same comments apply.

8. Complexity of PD Operators

We now turn to the computational complexity of PD operators. First we need to turn the computation of a PD operator into a decision problem.

We define a *PD revision instance* (or *PDR instance* for short) to be a tuple $\langle P, R, K, \varphi, \psi \rangle$ where,

- P is a nonempty set of propositional variables.
- R is a function $P \mapsto [1..|P|]$, represented as a set of ordered pairs (p, i) where $p \in P$ and $1 \leqslant i \leqslant |P|$.
- K is a consistent set of clauses over the variables in P.[22]
- φ is a consistent set of clauses over the variables in P.
- ψ is a consistent set of clauses over the variables in P.

A PDR instance $Q = \langle P, R, K, \varphi, \psi \rangle$ represents a specific belief revision scenario. In particular, P represents the set of propositional variables over which beliefs are expressed, K represents the (base of) the current belief set, φ is the sentence by which K is revised, and ψ is the sentence we wish to test at the revised state (see below). The function R is used to represent a preorder $\trianglelefteq$ over the variables in P; in particular, for any $p, q \in P$, $p \trianglelefteq q$ iff $R(p) \leqslant R(q)$. Clearly $\trianglelefteq$ generates PD preorders, which in turn define a PD revision operator $*$. The decision problem associated with the PDR instance Q, which we call the *PD revision problem*, is whether $Cn(K) * \varphi \models \psi$.

[22] A clause is defined as a (finite) disjunction of literals.

Observe that if $R(p) = 1$ for all $p \in P$, then $*$ reduces to Dalal's operator $\circ$. In [8], it was shown that deciding if $Cn(K) \circ \varphi \models \psi$ is $P^{NP[O(\log n)]}$-complete (see their Theorem 6.9).[23] Hence we immediately derive the following result.

Theorem 11. *The PD revision problem is* $P^{NP[O(\log n)]}$*-hard.*

An upper bound to the computational complexity of the PD revision problem is given by the following theorem:

Theorem 12. *The PD revision problem belongs to* $P^{NP[O(\sqrt{n}\log n)]}$.

Proof. Let $Q = \langle P, R, K, \varphi, \psi \rangle$ be a PDR instance and let $*$ be the PD revision function associated with Q. We prove membership in the class $P^{NP[O(\sqrt{n}\log n)]}$ by outlining an algorithm that decides $Cn(K) * \varphi \models \psi$ with $O(\sqrt{n}\log n)$ calls to an NP oracle, where $n = |P|$.

The algorithm has three phases. In the first phase we compute the smallest number k in the set $\{|Diff(w, r)|\colon w \in [K] \text{ and } r \in [\varphi]\}$. Observe that $k \leqslant n$. Hence we can use binary search to determine k in $\log n$ steps, where at each step the question of whether $k \leqslant j$ (i.e., whether there exist $w \in [K]$ and $r \in [\varphi]$ such that $|Diff(w, r)| \leqslant j$), is decided with a call to the NP oracle.[24]

Before proceeding with the second phase of the algorithm we need some further notation and terminology.

Let $P_1, P_2, \ldots P_m$ be the equivalence classes induced from $\trianglelefteq$ (alias R); i.e., the P_i's are nonempty, pairwise disjoint sets, such that their union equals P, and moreover, for

[23] We recall that a decision problem Π belongs to the class P^{NP} if it can be solved in polynomial time by a deterministic Turing machine M with an NP oracle. If in addition M can solve any instance of Π of length n, with no more than $g(n)$ calls to its NP oracle, we say that Π belongs to $P^{NP[g(n)]}$ – see the review in [8] for more details. For an excellent text on NP-completeness refer to [33].

[24] This first phase is identical to the first phase of the algorithm described in proof of Theorem 6.9 in [8]. As noted in [8], deciding whether $k \leqslant j$ is clearly in NP: a positive response can be verified by guessing two worlds w, r, and confirming that, firstly, $w \models K$, secondly, $r \models \varphi$, and thirdly $|Diff(w, r)| \leqslant j$. All this can be done in (non-deterministic) linear time.

any $p, q \in P$, $p \trianglelefteq q$ iff $p \in P_i$, $q \in P_j$ and $i \leqslant j$. Clearly, $P_1, \ldots P_m$ can be computed in deterministic polynomial time.

For a set of propositional variables $S \subseteq P$, define the *profile of S* to be the tuple $\langle j_1, \ldots, j_m \rangle$, where $j_1 = |S \cap P_1|$, …, $j_m = |S \cap P_m|$; that is, the profile of S is the number of elements that S shares with each equivalence class $P_1, \ldots, P_m$. A crucial observation, is that all $\trianglelefteq$-minimal elements in $\{Diff(w, r)$: $w \in [K]$ and $r \in [\varphi]\}$, have the same profile. We call this profile, the *minimal profile* wrt K and φ and we shall denote it by $\langle y_1, \ldots, y_m \rangle$. At the second phase our algorithm computes the numbers $y_1, \ldots, y_m$.

Define $x_1 = |P_1|$, …, $x_m = |P_m|$. Hence, $x_i > 0$ and $\sum_{i=1}^{m} x_i = n$. Moreover, $0 \leqslant y_i \leqslant x_i$, for all $1 \leqslant i \leqslant m$; also it is not hard to see that $\sum_{i=1}^{m} y_i = k$.

The second phase of the algorithm starts with the computation of y_1. Observe that y_1 is the maximal size of $Diff(w, r) \cap P_1$ under the constraints that $w \in [K]$, $r \in [\varphi]$ and $|Diff(w, r)| = k$. Hence y_1 can be computed with binary search in $log\, x_1$ steps, where at each step the question whether $y_1 \geqslant j$ (i.e., whether there exist $w \in [K]$ and $r \in [\varphi]$ such that $|Diff(w, r)| = k$ and $|Diff(w, r) \cap P_1| \geqslant j$), is decided with a call to the NP oracle.

Now the rest of the y_i's can be computed based on the following observation: y_{i+1} is the maximal size of $Diff(w, r) \cap P_{j+1}$, under the constraints that $w \in [K]$, $r \in [\varphi]$, $|Diff(w, r)| = k$, and $|Diff(w, r) \cap P_j| = y_j$ for all $1 \leqslant j \leqslant i$. Hence y_{i+1} can be computed with binary search in $log\, x_{i+1}$ steps, where at each step the question whether $y_{i+1} \geqslant j$ is decided with a call to the NP oracle.

The whole minimal profile $\langle y_1, \ldots, y_m \rangle$ can then be computed in polynomial time with $log\, x_1 + \ldots + log\, x_m$ calls to an NP oracle. Given that $\sum_{i=1}^{m} x_i = n$, from the inequality of arithmetic and geometric means we derive that $log\, x_1 + \ldots + log\, x_m \leqslant \frac{1}{2}\sqrt{n}\, log\, n$. Hence in the first two phases our algorithm makes at most $O(\sqrt{n}\, log\, n)$ calls to the NP oracle.

The third phase involves only one extra call to the oracle. In particular, the algorithm tests, with the aid of the NP oracle, whether there are worlds $w \in [K]$ and $r \in [\varphi]$

such that $Diff(w, r)$ has profile $\langle y_1, \ldots, y_m \rangle$ and moreover $r \models \neg\psi$. If the answer is positive, the algorithm returns "no" to the original question '$Cn(K) * \varphi \models \psi ?''$; otherwise it returns "yes". $\square$

Theorems 11, 12 show that the PD revision problem belongs to the second level of the polynomial hierarchy. This is the same level where (the computation of) Dalal's operator belongs. Hence the added expressivity of PD operators doesn't have any drastic effects in time complexity.

We conclude this section by considering the restriction of the PD revision problem to the special case of *Horn clauses*.

In particular, let $Q = \langle P, R, K, \varphi, \psi \rangle$ be a PDR instance such that all clauses in K, φ and ψ are Horn clauses.[25] We shall denote by $||K||$, $||\varphi||$, $||\psi||$, the size of K, φ, and ψ respectively. Eiter and Gottlob showed in [8] that in this case and for Dalal's operator $\square$, the query "$Cn(K) \square \varphi \models \psi ?''$ can be computed in $O(||K|| \cdot ||\psi||)$ time, provided that the size of φ is bounded by a constant.

This result can be extended, with some adjustments, to apply to any PD operator, using the same line of reasoning adopted in the proof of Theorem 8.3 in [8].

In particular, we show that when K, φ, and ψ are sets of Horn clauses, then for any PD operator $*$ the query "$Cn(K) \square \varphi \models \psi ?''$ can be computed in time $O(||K||^2 \cdot (log^2(||K||) + ||\psi||))$, provided that the size of φ is not larger than $log(||K||)$.

Hence, compared to Theorem 8.3 in [8], Theorem 13 below, firstly, relaxes the requirement on the size of φ (from bounding it by a constant to bounding it by $log(||K||)$), and secondly, it applies to any PD operator (not just to Dalal's).

On the other hand this generalisation comes at a price: the time complexity of deciding whether $Cn(K) \square \varphi \models \psi$, though still polynomial, is higher than that in Theorem 8.3 in [8] (yet see Corollary 3 at the end of this section).

We note that having that the size of revision input φ to be small compared to the

[25] We recall that a Horn clause is a disjunction of literals such that at at most one of them is an atom.

size of the initial knowledge base K (in particular, *logarithmically smaller*), is a natural assumption; the new input φ is typically the result of direct observation or feedback from a reliable source, which is nowhere near in size to the size of the initial belief corpus K.[26]

Theorem 13. *Let $\langle P, R, K, \varphi, \psi \rangle$ be a PDR instance and let $*$ be the PR revision operator induced by it at $Cn(K)$. Assume that K, φ, and ψ are sets of Horn clauses, and that $||\varphi|| \leqslant log(||K||)$. Then, deciding if $Cn(K) * \varphi \models \psi$ can be computed in $O(||K||^2 \cdot (log^2(||K||) + ||\psi||))$ time.*

Proof. As already mentioned, our proof follows the same line of reasoning as the proof of Theorem 8.3 in [8].

First some notation. For any literal x, we define the *variable of* x, denote x_v as follows: if $x \in P$ then $x_v = x$; if on the other hand $x = \neg y$ for some $y \in P$, then $x_v = y$. Moreover, for a set of literals A, by A_v we denote that set of variables of the literals in A; i.e., $A_v = \{x_v : x \in A\}$.

Let X be the set of all variables that appear (with or without negation) in φ. Since $||\varphi|| \leqslant log(||K||)$, it clearly follows that there are at most $log(||K||)$ variables in X.

Let $\trianglelefteq$ be the preorder over P induced from R and let w be any world in $[K]$. A crucial observation is that all worlds in $[\varphi]$ that differ minimally from w (wrt $\trianglelefteq$), agree with w over all variables outside X. To see this, consider any $r \in [\varphi]$ such that for some $q \in P - X$, $q \in \mathit{Diff}(w, r)$. Define r' to be the world that agrees with r in all variables except q. Clearly then, since $r \in [\varphi]$ and q does not appear in φ, we derive that $r' \in [\varphi]$. Moreover, by construction, $|\mathit{Diff}(w, r')| < |\mathit{Diff}(w, r)|$ and therefore, $\mathit{Diff}(w, r') \vartriangleleft \mathit{Diff}(w, r)$. Hence r doesn't differ minimally from w among the worlds in $[\varphi]$.

We shall denote by n the size of K. As noted earlier there are at most $log\, n$ atoms

[26] On the other hand, this assumption will prevent the use of a belief revision solver in the context of *belief merging*, [34]. We thank the anonymous reviewer for raising this point.

in X. Hence there are at most n sets of literals A such that $A_v = X$. We shall denote this set by $\mathcal{M}_X$; i.e. $\mathcal{M}_X = \{A : A$ is a set of literals such that $A_v = X\}$.

Since φ is Horn, we can check whether A is consistent with φ in time $O(log\, n)$, for any $A \in \mathcal{M}_X$. Notice that by construction, it holds that for any $A \in \mathcal{M}_X$, A is consistent with φ iff $A \models \varphi$. Hence we can construct the set of all elements of $\mathcal{M}_X$ that satisfy φ, denoted $[\varphi]_X$, in time $O(n\, log\, n)$.

Likewise, for any $B \in \mathcal{M}_X$, can check whether B is consistent with K in time $O(n)$. Hence we can construct the set of all elements of $\mathcal{M}_X$ that are consistent with K, denoted $[K]_X$, in time $O(n^2)$. Observe that by construction, for any $B \in \mathcal{M}_X$, $B \in [K]_X$, iff there is a $w \in [K]$ such that $B \subseteq w$.

For any $A, B \in \mathcal{M}_X$, let us denote by $\mathit{Diff}_X(A, B)$ the set of variables over which A and B disagree; i.e. $\mathit{Diff}_X(A, B) = \{x \in X : x \in A$ and $\neg x \in B\} \cup \{x \in X : \neg x \in A$ and $x \in B\}$. Moreover, by $[\varphi]_X^{min}$ we shall denote the set of all elements in $[\varphi]_X$ that differ *minimally wrt* $\trianglelefteq$ from some $B \in [K]_X$; i.e. $[\varphi]_X^{min} = \{A \in [\varphi]_X$: for some $B \in [K]_X$, $\mathit{Diff}_X(A, B) \trianglelefteq \mathit{Diff}_X(A', B')$, for all $A' \in [\varphi]_X$ and $B' \in [K]_X\}$.

Clearly, for each $A \in [\varphi]_X$ and $B \in [K]_X$, $\mathit{Diff}_X(A, B)$ can compute in time $O(log^2 n)$. Moreover, it is not hard to see that for any $S, S' \subseteq X$ to compute whether $S \trianglelefteq S'$ never takes longer than $O(log^2 n)$ time. Hence, $[\varphi]_X^{min}$ can be computed in time $O(n^2\, log^2 n)$.

In the next step we shall use K and $[\varphi]_X^{min}$ to construct a sentence K' such that $Cn(K') = Cn(K) * \varphi$.

For this step we adopt the procedure described in the proof of Theorem 8.3 in [8].

In particular, for every element $A \in [\varphi]_X^{min}$ we shall construct a conjunction of Horn clauses D_A; K' will then be defined as the disjunction of all D_A, such that $A \in [\varphi]_X^{min}$.

For any $A \in [\varphi]_X^{min}$, define K_A to be the set of Horn clauses that is produced from K by replacing every variable in X with the truth value T (true) if $x \in A$, and with F (false) otherwise. K_A can of course be simplified by removing all clauses containing the truth value T (or containing $\neg F$), and deleting the truth value F from the remaining clauses. Clearly, (the simplified) K_A is a set of Horn clauses. D_A is then defined as the conjunction of all clauses in (the simplified) K_A with A. Since A is a conjunction of literals, D_A is a Horn formula (i.e. a conjunction of Horn clauses). Moreover, from the construction of D_A, is it clear that it can be computed in time $O(n\,log\,n)$.

Hence K', defined as the disjunction of all D_A where $A \in [\varphi]_X^{min}$, can be computed in time $O(n^2\,log\,n)$.

From the construction of K', it is not hard to see that it is indeed logically equivalent to $Cn(K) * \varphi$; i.e. $Cn(K') = Cn(K) * \varphi$.

Therefore $Cn(K) * \varphi \models \psi$ is equivalent to $K' \models \psi$. Moreover, since $K' = \bigvee\{D_A : A \in [\varphi]_X^{min}\}$, $K' \not\models \psi$, iff there is an $A \in [\varphi]_X^{min}$, such that $D_A \not\models \psi$. Given that both D_A and ψ are Horn formulas, and moreover, by construction, the size of D_A is at most $n + log\,n$, it follows that we can compute whether $D_A \not\models \psi$ in time $O(n \cdot ||\psi||)$. Then we can compute whether $K' \models \psi$ in time $O(n^2 \cdot ||\psi||)$.

Summarising the complexity of the steps involved in computing the answer to the query $Cn(K) * \varphi \models \psi$, we need at most $O(n\,log\,n)$ time to compute $[\varphi]_X$, $O(n^2)$ time to compute $[K]_X$, $O(n^2\,log^2 n)$ time to compute $[\varphi]_X^{min}$, $O(n^2 log\,n)$ time to compute K', and $O(n^2 \cdot ||\psi||)$ to compute whether $K' \models \psi$. Since these steps occur sequentially, the time complexity of the overall procedure is at most $O(n^2 \cdot (log^2 n + ||\psi||))$, or equivalently $O(||K||^2 \cdot (log^2(||K||) + ||\psi||))$ $\square$

We note that if $||\varphi||$ is bounded by a constant, then by following the exact same steps as in the proof above, we derive the corollary below:

Corollary 3. *Let $\langle P, R, K, \varphi, \psi \rangle$ PDR instance and let $*$ be the PR revision operator induced by it at $Cn(K)$. Assume that K, φ, and ψ are sets of Horn clauses, and that $||\varphi|| \leqslant k$ for some constant k. Then, deciding if $Cn(K) * \varphi \models \psi$ can be computed in $O(||K|| \cdot ||\psi||)$ time.*

Clearly, from the above corollary it follows that if the size of ψ is also bounded by a constant, then deciding if $Cn(K) * \varphi \models \psi$ can be computed in time *linear* to the size of K.

9. Conclusion

To deal with the problem of the high representational cost required by any general AGM belief revision solver, in this paper we introduced and studied a new family of concrete AGM revision operators called PD operators.

Any PD operator can be constructed from a single preorder over atoms. Moreover PD operators are expressive enough to encode challenging belief revision scenarios discussed in the literature.

In addition to the semantic definition of PD operator (essentially a generalisation of Dalal's approach), in this article we provided an axiomatic characterisation of PD operators. Moreover a number of attractive properties of PD operators were established, including their compliance with Parikh's notion of relevance-sensitive belief revision.

Finally we have studied the computational complexity of PD operators showing that they lie at the same level of the polynomial hierarchy as Dalal's operator (despite the extra expressivity). In the special case of Horn formulas, PD operators become tractable provided that the size of the new evidence is small compared to that of the initial knowledge base. The complexity further reduces to *linear time* to the size of the initial (Horn) knowledge base, if the size of the revision queries is bounded by a constant.

A different direction for tackling the high complexity of belief revision has been proposed by Pfandler, et. al, in [35]. Although the authors focus on Satoh's operator, [12], which does not satisfy the AGM postulates for belief revision, their techniques may nevertheless be applicable to PD Operators. This is a interesting avenue for future work.

Another promising direction for future work future work is to investigate the possibility of *evolving* the preorder $\trianglelefteq$ over atoms that defines a PD operator, in response to the new information received.[27] This ought to be done with no or very little addi-

[27] In this paper we have assumed that $\trianglelefteq$ is not affected by new evidence.

tional representational and computational cost, for otherwise the benefits of using PD operators would be cancelled.

Acknowledgements

The authors would like to thank Theofanis Aravanis for his comments on this work. The anonymous referees provided detailed and valuable remarks that helped us improve our paper, and for which we are grateful. This research was supported by the Federal Ministry of Education and Research (BMBF), Germany under the project LeibnizKI-Labor with grant No. 01DD20003, and by the University of Huddersfield.

References

[1] P. Peppas, M.-A. Williams, Parametrised difference revision, in: Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR 2018), 2018.

[2] P. Peppas, M.-A. Williams, Kinetic consistency and relevance in belief revision, in: Proceedings of the 15th European Conference on Logics in Artificial Intelligence (JELIA 2016), 2016.

[3] C. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet functions for contraction and revision, Journal of Symbolic Logic 50 (1985) 510–530.

[4] P. Peppas, Belief revision, in: F. van Harmelen, V. Lifschitz, B. Porter (Eds.), Handbook of Knowledge Representation, Elsevier Science, 2008, pp. 317–359.

[5] T. Chou, M. Winslett, The implementation of a model-based belief revision system, ACM SIGART Bulletin 2 (3) (1991) 28–34.

[6] M.-A. Williams, A. Sims, Saten: An object-oriented web-based revision and extraction engine, in: International Workshop on Nonmonotonic Reasoning (NMR-2000). Online Computer Science Abstract, http://arxiv.org/abs/cs.AI/0003059/, 2000.

[7] C. Beierle, G. Kern-Isberner, A verified asml implementation of belief revision, Lecture Notes in Computer Science 5238 (2008) 98–111.

[8] T. Eiter, G. Gottlob, On the complexity of propositional knowledge base revision, updates, and counterfactuals, Artificial Intelligence 57 (1992) 227–270.

[9] H. Katsuno, A. Mendelzon, Propositional knowledge base revision and minimal change, Artificial Intelligence 52 (3) (1991) 263–294.

[10] A. Borgida, Language features and flexible handling of exceptions in information systems, ACM Transactions on Database Systems (TODS) 10 (4) (1985) 563–603.

[11] M. Winslett, Reasoning about action using a possible models approach, in: Proceedings of the 7th National Conference of the American Association for Artificial Intelligence (AAAI'88), 1988, pp. 89–93.

[12] K. Satoh, Nonmonotonic reasoning by minimal belief revision, in: Proceedings of the International Conference on Fifth Generation Computer Systems, Springer-Verlag, Tokyo, 1988, pp. 455–462.

[13] A. Weber, Updating propositional formulas, Proc. First Conference on Expert Database Systems (1986) 487–500.

[14] M. Dalal, Investigations into theory of knowledge base revision: Preliminary report, in: Proceedings of 7th National Conference of the American Association for Artificial Intelligence (AAAI'88), 1988, pp. 475–479.

[15] R. Parikh, Beliefs, belief revision, and splitting languages, in: Logic, Language, and Computation - CSLI Lecture Notes, Vol. 2, CSLI Publications, 1999, pp. 266–278.

[16] P. Gärdenfors, Knowledge in Flux, MIT press, 1988.

[17] A. Darwiche, J. Pearl, On the logic of iterated belief revision, Artificial Intelligence 89 (1997) 1–29.

[18] Y. Jin, M. Thielscher, Iterated belief revision, revised, Artificial Intelligence 171 (2007) 1–18.

[19] S. Konieczny, R. P. Perez, A framework for iterated revision, Journal of Applied Non-Classical Logics 10 (2000) 339–367.

[20] A. Nayak, M. Pagnucco, P. Peppas, Dynamic belief revision operators, Artificial Intelligence 146 (2003) 193–228.

[21] S. Konieczny, J. Lang, P. Marquis, DA^2 merging operators, Artificial Intelligence 157 (2004) 49–79.

[22] P. Peppas, M.-A. Williams, S. Chopra, N. Foo, Relevance in belief revision, Artificial Intelligence 229 (2015) 126–138.

[23] T. Aravanis, An ASP-based solver for parametrized-difference revision, Journal of Logic and Computation 32 (2022) 630–666.

[24] A. Hunter, J. Agapeyev, Epistemic entrenchment characterization of Parikh's axiom, in: AI 2019: Advances in Artificial Intelligence: 32nd Australasian Joint Conference, Adelaide, Australia, 2019.

[25] C. Areces, V. Becher, Iterable agm functions, Frontiers in Belief Revision, Applied Logic Series 22 (2001) 239–270.

[26] T. Aravanis, On uniform belief revision, Journal of Logic and Computation 30 (2020) 1357–1376.

[27] T. Aravanis, P. Peppas, Theory-relational belief revision, Annals of Mathematics and Artificial Intelligence 90 (2022) 573–594.

[28] P. Heltweg, Implementing a structured approach to belief revision by deterministic switching between total preorders, Master Thesis, University of Hagen,Germany.

[29] B. Nebel, How hard is it to revise a belief base?, in: D. Dubois, H. Prade (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 3: Belief Change, Kluwer Academic, 1998, pp. 77–145.

[30] H. Rott, Shifting priorities: Simple representations for twenty-seven iterated theory change operators, Towards mathematical philosophy. Trends in logic 28 (2009) 269–296.

[31] M.-A. Williams, On the logic of theory base change, in: D.Pearce, L. Pereira (Eds.), European Conference on Logics in Artificial Intelligence (JELIA'94), Springer-Verlag, 1994, pp. 86–105.

[32] B. Nebel, Base revision operations and schemes: Semantics, representation, and complexity, in: Proceedings of the 11th European Conference on Artificial Intelligence, Wiley, 1994, pp. 341–345.

[33] M. Garey, D. Johnson, Computers and Intractability – A Guide to the Theory of NP-Completeness, W.H. Freeman, 1979.

[34] S. Konieczny, R. P. Pérez, Logic based merging, Journal of Philosophical Logic 40 (2011) 239–270.

[35] A. Pfandler, S. Rümmele, J. P. Wallner, S. Woltran, On the parameterized complexity of belief revision, in: Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI-15), Buenos Aires, Argentina, 2015.