

Countermodel construction via optimal hypersequent calculi for non-normal modal logics^{*}

Tiziano Dalmonte¹, Björn Lellmann², Nicola Olivetti¹, and Elaine Pimentel³

¹ Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France - tiziano.dalmonte@etu.univ-amu.fr, nicola.olivetti@lis-lab.fr

² Technische Universität Wien, Austria - lellmann@logic.at

³ UFRN Natal, Brazil - elaine.pimentel@gmail.com

Abstract. We develop semantically-oriented calculi for the cube of non-normal modal logics and some deontic extensions. The calculi manipulate hypersequents and have a simple semantic interpretation. Their main feature is that they allow for direct countermodel extraction. Moreover they provide an optimal decision procedure for the respective logics. They also enjoy standard proof-theoretical properties, such as a syntactical proof of cut-admissibility.

Keywords: Hypersequents, Non-normal modal logics, Countermodel construction, Cut elimination.

1 Introduction

Non-normal modal logics–NNMLs for short–have a long history, going back to the seminal works by Kripke, Montague, Segeberg, Scott, and Chellas (see [3] for an introduction). They are “non-normal” as they do not contain all axioms of minimal *normal* modal logic **K**. NNMLs find an interest in several areas: in epistemic reasoning they offer a simple (although partial) solution to the problem of logical omniscience (see [19]); in deontic logic, they allow avoiding well-known paradoxes (such as Ross’s Paradox) and to represent conflicting obligations (see [8]); NNMLs are needed also when $\Box A$ is interpreted as “ A is true in most of the cases” [1]; finally NNMLs naturally arise in game-theoretical interpretation of $\Box A$ as “the agent has a winning strategy to bring about A ” (indeed, non-normal monotonic logic **M** can be seen as a 2-agent case of coalition logic with determinacy [18]). In this work, we consider the *classical cube* on NNMLs, given by the extensions of the minimal modal logic **E**, containing only the *congruence rule*, with axioms C, M and N.

NNMLs have a well-understood semantics defined in terms of neighbourhood models [16]: in these models each world w has an associated set of neighbourhoods $\mathcal{N}(w)$, each one of them being a set of worlds/states. If we accept the

^{*} Supported by WWTF project MA16-28 and by Project TICAMORE ANR-16-CE91-0002-01.

traditional interpretation of a “proposition” as a set of worlds (= its truth set), we can think of each neighbourhood in $\mathcal{N}(w)$ as a proposition: a formula $\Box A$ is true in a world w if “the proposition” A , *i.e.* the truth-set of A , belongs to $\mathcal{N}(w)$. The classical cube can be modelled by imposing additional closure properties of the set of neighbourhoods. In this work we adopt a variant of neighbourhood semantics defined in terms of bi-neighbourhood models [4]: in these structures each world has associated a set of *pairs* of neighbourhoods. The intuition is that the two components of a pair provide positive and negative support for a modal formula, being more natural for “non-monotonic” logics (*i.e.* not containing axiom M). The reason is that, instead of specifying exactly the truth sets in $\mathcal{N}(w)$, the pairs of neighbourhoods specify just lower and upper bounds of truth sets, so that the same pair may be a “witness” for several propositions. This makes the generation of *countermodels* easier, as shown in this paper.

It is curious to note that, although some proof-systems for NNMLs have been proposed in the past, countermodel extraction has been rarely addressed and complexity is seldom analysed. Indeed, the works [11, 7, 14, 4] propose countermodel extraction, but all of them require either a complicated procedure or an extended language with labels. [12] presents a nested sequent calculus for a logic combining normal and monotone non-normal modal logic that supports countermodel extraction, but the nested sequent structure is not suitable for logics lacking monotonicity. On the other hand, cut-free sequent/linear nested calculi for the classical cube and its extensions with standard axioms of normal modal logics (the non-normal counterpart of logics from **K** to **S5**) are studied in [9, 10, 13, 15]; however neither semantic completeness and countermodel extraction, nor complexity are studied.

In this work, we intend to fill this gap by proposing *modular* calculi for the classical cube (and also some deontic extensions) that provide *direct countermodel extraction* and are of *optimal complexity*. Our calculi are semantically based on the bi-neighbourhood models, and have two syntactic features: they manipulate hypersequents and sequents may contain blocks of \Box -ed formulas in the antecedent. A hypersequent [2] is just a multiset of sequents and can be understood as a (meta-logical) disjunction of sequents. Sequents within hypersequents can be read as formulas of the logic. Blocks of formulas are interpreted as conjunction of positive \Box -ed formulas. Intuitively each block represents a neighbourhood satisfying one or more \Box -ed formulas, and this allows for the formulation of modular calculi for the whole cube. The advantage of using hypersequents is that all rules become invertible, thus there is no need for backtracking in proof search. For the same reason, the hypersequent calculi provide directly countermodel extraction: from *one* failed proof, it can be extracted *directly* a countermodel in the bi-neighbourhood semantics of the sequent/formula at the root of the derivation. For logics without the C axiom, our calculi also provide an *optimal* NP/CoNP decision procedure. For logics including C, we can still obtain an optimal PSPACE decision procedure but, as usual in modal logic, at the cost of sacrificing the invertibility of some rules. Finally, the calculi have

also good proof-theoretical properties, as they support a syntactic proof of cut admissibility.

It turns out that our calculi can be modularly extended by adding modal axioms. We illustrate this by extending them with axioms T, P, D, the last two are of interest in deontic logic.

All in all, we believe that the structure of our calculi, namely hypersequents with blocks, is adequate for NNMLs from a semantical, computational and a proof-theoretical point of view since it: (i) has a semantic interpretation; (ii) allows direct countermodel generation; (iii) has optimal complexity; and (iv) has good proof-theoretical properties.

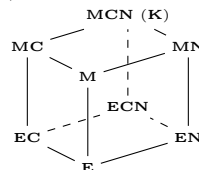
2 Non-normal modal logics

In this section, we present the classical cube of NNMLs, both axiomatically and semantically in terms of neighbourhood models. We also present bi-neighbourhood models, a variant of the neighbourhood semantics introduced in [4]. The propositional language \mathcal{L} contains *formulas* given by the following grammar: $A ::= p \mid \top \mid \perp \mid A \rightarrow A \mid \Box A$, where $p \in \text{Atm}$, the set of propositional variable symbols. Other propositional connectives are defined by the standard equivalences. The minimal logic **E** in the language \mathcal{L} is given Hilbert-style by extending classical propositional logic with only the *congruence rule*

$$\text{RE} \frac{A \rightarrow B \quad B \rightarrow A}{\Box A \rightarrow \Box B}$$

The *classical cube* (below on the right) is formed by extending **E** with any combination of axioms M, C, and N (below on the left).

$$\begin{array}{ll} \text{M} & \Box(A \wedge B) \rightarrow \Box A \\ \text{C} & \Box A \wedge \Box B \rightarrow \Box(A \wedge B) \\ \text{N} & \Box \top \end{array}$$



As usual for logics containing M we omit E, *e.g.*, we write **MC** for **EMC**. In the following, for any system **L** of the cube we denote with **L**^{*} any extension of **L** obtained by adding one or more of these axioms. We recall that axioms M and N are, respectively, equivalent to the rules RM ($A \rightarrow B / \Box A \rightarrow \Box B$) and RN ($A / \Box A$), and that axiom K ($\Box(A \rightarrow B) \rightarrow \Box A \rightarrow \Box B$) is derivable from M and C. As a consequence, the top system **MCN** is equivalent to **K**, the weakest normal modal logic.

The standard semantics for NNMLs is defined in terms of so-called neighbourhood (or minimal) models [3].

Definition 1. A standard neighbourhood model is a tuple $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}_s, \mathcal{V} \rangle$, where \mathcal{W} is a non-empty set, \mathcal{V} is a valuation function, and \mathcal{N}_s is a neighbourhood function $\mathcal{W} \rightarrow \mathcal{P}(\mathcal{P}(\mathcal{W}))$. A model \mathcal{M} is supplemented if $\alpha \in \mathcal{N}_s(w)$ and $\alpha \subseteq \beta \subseteq \mathcal{W}$ implies $\beta \in \mathcal{N}_s(w)$, it contains the unit if $\mathcal{W} \in \mathcal{N}_s(w)$ for all $w \in \mathcal{W}$,

and it is closed under intersection if $\alpha, \beta \in \mathcal{N}_s(w)$ implies $\alpha \cap \beta \in \mathcal{N}_s(w)$. The standard forcing relation for boxed formulas is: $\mathcal{M}, w \Vdash_s \Box A$ iff $[A]_{\mathcal{M}} \in \mathcal{N}_s(w)$, where $[A]_{\mathcal{M}}$ denotes the set $\{v \in \mathcal{W} \mid \mathcal{M}, v \Vdash_s A\}$ of the worlds v that force A , also called the truth set of A .

Neighbourhood models characterise modularly the classical cube of NNMLs [3] in the sense that a formula A is a theorem of **E** if and only if it is valid in all neighbourhood models. Furthermore, A is a theorem of **E+(M/C/N)** iff it is valid, respectively, in all models that are supplemented (M), closed under intersection (C), and contain the unit (N) (and any combination of these previous axioms/conditions).

Here we consider bi-neighbourhood semantics [4], a variant of neighbourhood semantics. In this semantics, neighbourhoods come in pairs, the two components provide, so to say, “positive” and “negative” support for a modal formula.

Definition 2. A bi-neighbourhood model is a tuple $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}_b, \mathcal{V} \rangle$, where \mathcal{W} is a non-empty set, \mathcal{V} is a valuation function, and \mathcal{N}_b is a bi-neighbourhood function $\mathcal{W} \rightarrow \mathcal{P}(\mathcal{P}(\mathcal{W}) \times \mathcal{P}(\mathcal{W}))$. We say that \mathcal{M} is a M-model if $(\alpha, \beta) \in \mathcal{N}_b(w)$ implies $\beta = \emptyset$, it is a N-model if for all $w \in \mathcal{W}$ there is $\alpha \subseteq \mathcal{W}$ such that $(\alpha, \emptyset) \in \mathcal{N}_b(w)$,⁴ and it is a C-model if $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathcal{N}_b(w)$ implies $(\alpha_1 \cap \alpha_2, \beta_1 \cup \beta_2) \in \mathcal{N}_b(w)$. The forcing relation for boxed formulas is as follows:

$$\mathcal{M}, w \Vdash_b \Box A \quad \text{iff} \quad \text{there is } (\alpha, \beta) \in \mathcal{N}_b(w) \text{ s.t. } \alpha \subseteq [A]_{\mathcal{M}} \text{ and } \beta \subseteq [\neg A]_{\mathcal{M}}.$$

Every standard model gives rise to a bi-neighbourhood model, by taking for each neighbourhood $\alpha \in \mathcal{N}_s(w)$, the pair $(\alpha, \mathcal{W} \setminus \alpha)$. Conversely, every bi-neighbourhood model can be transformed into a standard model by assigning to $\mathcal{N}_s(w)$, for each pair $(\alpha, \beta) \in \mathcal{N}_b(w)$, the subsets γ such that $\alpha \subseteq \gamma \subseteq \mathcal{W} \setminus \beta$. In this sense α and β give upper and lower bounds for neighbourhoods in standard models. For the non-monotonic case there is also a finer transformation which depends on a set \mathcal{S} of formulas [4].

Proposition 1. Let $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}_b, \mathcal{V} \rangle$ be a bi-neighbourhood model and \mathcal{S} be a set of formulas closed under subformulas. We define the standard neighbourhood model $\mathcal{M}' = \langle \mathcal{W}', \mathcal{N}_s, \mathcal{V}' \rangle$ by taking $\mathcal{W}' = \mathcal{W}$, $\mathcal{V}' = \mathcal{V}$ and for all $w \in \mathcal{W}$, $\mathcal{N}_s(w) = \{[A]_{\mathcal{M}} \mid w \Vdash_b \Box A \text{ and } \Box A \in \mathcal{S}\}$. Then for all $w \in \mathcal{W}$, $A \in \mathcal{S}$,

$$\mathcal{M}', w \Vdash_s A \quad \text{iff} \quad \mathcal{M}, w \Vdash_b A.$$

Further, if \mathcal{M} is a N-model and $\Box \top \in \mathcal{S}$, then \mathcal{M}' contains the unit, and if \mathcal{M} is a C-model and $\Box A, \Box B \in \mathcal{S}$ implies $\Box(A \wedge B) \in \mathcal{S}$, then \mathcal{M}' is closed under intersection.

The transformation in Prop. 1 produces models with a smaller neighbourhood function. Whereas in models produced by the first transformation the size of $\mathcal{N}_s(w)$ can be exponential with respect to $\mathcal{N}_b(w)$, here the size of $\mathcal{N}_s(w)$ is linearly bounded by the number of boxed formulas in \mathcal{S} . As a paradigmatic case,

⁴ The N-model condition in [4] was slightly different. However, it is easy to verify that the two conditions are equivalent with respect to the validity of formulas.

suppose there is $(\emptyset, \emptyset) \in \mathcal{N}_b(w)$. By the rough transformation \mathcal{N}_s would contain all subsets of \mathcal{W} , whereas by the finer one it would contain only the truth sets of some boxed formulas.

While the two semantics characterise equally well the classical cube, we shall see in Section 5 that bi-neighbourhood semantics is more suited for direct countermodels extraction from failed proofs.

3 Hypersequent calculi

We now move our attention to proof theory. We will construct our calculi in the *hypersequent* framework (see, e.g., [2]). This choice is motivated mainly by the possibility of getting direct countermodel extraction, as detailed in the following. Moreover, our calculi will contain additional structures, called *blocks*, which are used to collect boxed formulas.

Our calculi are built as follows: A *block* is a structure $\langle \Sigma \rangle$, where Σ is a multiset of formulas of \mathcal{L} . A *sequent* is a pair $\Gamma \Rightarrow \Delta$, where Γ is a multiset of formulas and blocks, and Δ is a multiset of formulas. We sometimes consider $\text{set}(\Gamma)$: the *support* of a multiset Γ , i.e. the set of its elements disregarding multiplicities. A *hypersequent* is a multiset $S_1 \mid \dots \mid S_n$, where S_1, \dots, S_n are sequents. S_1, \dots, S_n are called *components* of the hypersequent. Single sequents can be interpreted as formulas of the logic in the following manner:

$$i(A_1, \dots, A_n, \langle \Sigma_1 \rangle, \dots, \langle \Sigma_m \rangle \Rightarrow B_1, \dots, B_k) = \bigwedge_{i \leq n} A_i \wedge \bigwedge_{j \leq m} \square \bigwedge \Sigma_j \rightarrow \bigvee_{\ell \leq k} B_\ell.$$

We say that a sequent S is *valid* in a (bi-)neighbourhood model \mathcal{M} ($\mathcal{M} \models S$) if for all $w \in \mathcal{M}$, $\mathcal{M}, w \Vdash i(S)$; and that a hypersequent H is *valid* in \mathcal{M} ($\mathcal{M} \models H$) if $\mathcal{M} \models S$ for some $S \in H$. Finally, we say that H is valid in (M/C/N-)models if it is valid in all models of that kind.

The hypersequent calculi \mathcal{H}_{E^*} are defined by the rules in Fig. 1, in particular: $\mathcal{H}_E :=$ propositional rules + \square_L + \square_R ; $\mathcal{H}_{EN} := \mathcal{H}_E + \mathbf{N}$; $\mathcal{H}_{EC} := \mathcal{H}_E + \mathbf{C}$; $\mathcal{H}_{ECN} := \mathcal{H}_E + \mathbf{C} + \mathbf{N}$; $\mathcal{H}_M :=$ propositional rules + \square_L + ${}_M \square_R$; $\mathcal{H}_{MN} := \mathcal{H}_M + \mathbf{N}$; $\mathcal{H}_{MC} := \mathcal{H}_M + \mathbf{C}$; and $\mathcal{H}_{MCN} := \mathcal{H}_M + \mathbf{C} + \mathbf{N}$.

Rules are given in their *kleene'd* versions, i.e., where the principal formula (or structure) is copied into every premiss. As usual, initial sequents *init* are restricted to propositional variables, but it is easy to see that $G \mid A, \Gamma \Rightarrow \Delta, A$ is derivable for any A . Note that the only rule which expands blocks is \mathbf{C} , then in absence of this rule the blocks occurring in a proof for a single formula contain only one formula. Examples of derivations are the following.

Example 1. Axioms \mathbf{M} , \mathbf{N} , \mathbf{C} are derivable in \mathcal{H}_{M^*} , \mathcal{H}_{EN^*} , and \mathcal{H}_{EC^*} , respectively

$$\frac{\frac{\langle A \wedge B \rangle, \square(A \wedge B) \Rightarrow \square A \mid A \wedge B \Rightarrow A}{} \quad \square_L}{\langle A \wedge B \rangle, \square(A \wedge B) \Rightarrow \square A} \quad \square_R}{\square(A \wedge B) \Rightarrow \square A} \quad \mathbf{M} \quad \frac{\langle \top \rangle \Rightarrow \square \top \mid \top \Rightarrow \top \quad \dots \mid \top \Rightarrow \top}{\langle \top \rangle \Rightarrow \square \top} \quad \square_R}{\Rightarrow \square \top} \quad \mathbf{N}$$

$$\begin{array}{c}
\hline
\text{init} \frac{}{G \mid p, \Gamma \Rightarrow \Delta, p} \quad \perp_{\text{L}} \frac{}{G \mid \perp, \Gamma \Rightarrow \Delta} \quad \top_{\text{R}} \frac{}{G \mid \Gamma \Rightarrow \Delta, \top} \\
\rightarrow_{\text{L}} \frac{G \mid A \rightarrow B, \Gamma \Rightarrow \Delta, A \quad G \mid B, A \rightarrow B, \Gamma \Rightarrow \Delta}{G \mid A \rightarrow B, \Gamma \Rightarrow \Delta} \quad \rightarrow_{\text{R}} \frac{G \mid A, \Gamma \Rightarrow \Delta, A \rightarrow B, B}{G \mid \Gamma \Rightarrow \Delta, A \rightarrow B} \\
\wedge_{\text{L}} \frac{G \mid A, B, A \wedge B, \Gamma \Rightarrow \Delta}{G \mid A \wedge B, \Gamma \Rightarrow \Delta} \quad \wedge_{\text{R}} \frac{G \mid \Gamma \Rightarrow \Delta, A \wedge B, A \quad G \mid \Gamma \Rightarrow \Delta, A \wedge B, B}{G \mid \Gamma \Rightarrow \Delta, A \wedge B} \\
\hline
\Box_{\text{L}} \frac{G \mid \langle A \rangle, \Box A, \Gamma \Rightarrow \Delta}{G \mid \Box A, \Gamma \Rightarrow \Delta} \quad \text{M}\Box_{\text{R}} \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid \Sigma \Rightarrow B}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B} \\
\Box_{\text{R}} \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid \Sigma \Rightarrow B \quad \{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid B \Rightarrow A\}_{A \in \Sigma}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B} \\
\text{N} \frac{G \mid \langle \top \rangle, \Gamma \Rightarrow \Delta}{G \mid \Gamma \Rightarrow \Delta} \quad \text{C} \frac{G \mid \langle \Sigma, \Pi \rangle, \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta} \\
\hline
\end{array}$$

Fig. 1: Rules of \mathcal{H}_{E^*} .

$$\frac{\frac{\langle A, B \rangle, \dots \mid A, B \Rightarrow A \wedge B \quad \langle A, B \rangle, \dots \mid A \wedge B \Rightarrow A \quad \langle A, B \rangle, \dots \mid A \wedge B \Rightarrow B}{\langle A, B \rangle, \langle A \rangle, \langle B \rangle, \Box A, \Box B, \Box A \wedge \Box B \Rightarrow \Box(A \wedge B)} \Box_{\text{R}}}{\frac{\langle A \rangle, \langle B \rangle, \Box A, \Box B, \Box A \wedge \Box B \Rightarrow \Box(A \wedge B)}{\Box A, \Box B, \Box A \wedge \Box B \Rightarrow \Box(A \wedge B)} \Box_{\text{L} \times 2}} \text{C} \\
\frac{\Box A, \Box B, \Box A \wedge \Box B \Rightarrow \Box(A \wedge B)}{\Box A \wedge \Box B \Rightarrow \Box(A \wedge B)} \wedge_{\text{L}}$$

We then have the following soundness theorem.

Theorem 1. *If H is derivable in $\mathcal{H}_{E(M/C/N)}$, then it is valid in $(M/C/N)$ -models.*

Proof. We show that whenever the premisses of a rule are valid, so is the conclusion. For propositional rules the proof is standard.

\Box_{L}) Assume $G \mid \langle A \rangle, \Gamma \Rightarrow \Delta$ valid. Then for all \mathcal{M} , $\mathcal{M} \models S$ for some $S \in G$, or $\mathcal{M} \models \langle A \rangle, \Box A, \Gamma \Rightarrow \Delta$. In the second case, for all $w \in \mathcal{M}$, $w \Vdash i(\langle A \rangle, \Box A, \Gamma \Rightarrow \Delta) = i(\Box A, \Box A, \Gamma \Rightarrow \Delta)$, which is equivalent to $i(\Box A, \Gamma \Rightarrow \Delta)$. Then $G \mid \Box A, \Gamma \Rightarrow \Delta$ is valid.

\Box_{R}) Let $\Sigma = C_1, \dots, C_n$ and assume $G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid \Sigma \Rightarrow B$ and $G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid B \Rightarrow C_i$ valid for all $1 \leq i \leq n$. Then for all \mathcal{M} , $\mathcal{M} \models S$ for some $S \in G$, or $\mathcal{M} \models \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B$. Otherwise $\Sigma \Rightarrow B, B \Rightarrow C_1, \dots, B \Rightarrow C_n$ are valid in \mathcal{M} . In the last case, for all $w \in \mathcal{M}$, $w \Vdash (\bigwedge \Sigma \rightarrow B) \wedge (B \rightarrow C_1) \wedge \dots \wedge (B \rightarrow C_n)$, that is $\mathcal{M} \models \bigwedge \Sigma \leftrightarrow B$. Then $\mathcal{M} \models \Box \bigwedge \Sigma \leftrightarrow \Box B$, so $\mathcal{M} \models \langle \Sigma \rangle \Rightarrow \Box B$. Therefore $G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B$ is valid.

$\text{M}\Box_{\text{R}}$) Analogous to \Box_{R} , by considering that in M-models $\mathcal{M} \models \bigwedge \Sigma \rightarrow B$ implies $\mathcal{M} \models \Box \bigwedge \Sigma \rightarrow \Box B$.

N) Assume $G \mid \langle \top \rangle, \Gamma \Rightarrow \Delta$ valid. Then for all \mathcal{M} , $\mathcal{M} \models S$ for some $S \in G$, or $\mathcal{M} \models \langle \top \rangle, \Gamma \Rightarrow \Delta$. In the second case, for all $w \in \mathcal{M}$, $w \Vdash i(\langle \top \rangle, \Gamma \Rightarrow \Delta)$, which is equivalent to $\Box \top \rightarrow i(\Gamma \Rightarrow \Delta)$. Since $\Box \top$ is valid in N-models, $\mathcal{M} \models \Gamma \Rightarrow \Delta$.

C) Assume $G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \langle \Sigma, \Pi \rangle, \Gamma \Rightarrow \Delta$ valid. Then for all \mathcal{M} , $\mathcal{M} \models S$ for some $S \in G$, or $\mathcal{M} \models \langle \Sigma \rangle, \langle \Pi \rangle, \langle \Sigma, \Pi \rangle, \Gamma \Rightarrow \Delta$. In the second case, for all $w \in \mathcal{M}$, $w \Vdash (\Box \wedge \Sigma \wedge \Box \wedge \Pi \wedge \Box (\wedge \Sigma \wedge \wedge \Pi)) \rightarrow i(\Gamma \Rightarrow \Delta)$. Since $\Box \wedge \Sigma \wedge \Box \wedge \Pi \rightarrow \Box (\wedge \Sigma \wedge \wedge \Pi)$ is valid in C-models, $w \Vdash (\Box \wedge \Sigma \wedge \Box \wedge \Pi) \rightarrow i(\Gamma \Rightarrow \Delta)$. Then $\mathcal{M} \models \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta$. \square

We make clear that to the purpose of having sound and complete calculi for NNMLs the hypersequent framework is not necessary, as for instance the sequent calculi in [11] show. Moreover, for the calculi \mathcal{H}_{E^*} whenever a hypersequent is derivable there is a component which is derivable. But as we shall see, the hypersequent framework is very adequate to extract countermodels from a single failed proof, ensuring at the same time good computational and structural properties. As a matter of fact, even in the bi-neighbourhood semantics, non-normal modal logics without monotonicity ultimately need to consider truth sets of formulas. Hence, to make our calculi suitable for a reasonably straightforward countermodel construction, we need to be able to represent essentially all worlds of a possible model in the data structure used by the calculus. While this could also be accomplished by, *e.g.* nested sequents, for obtaining small countermodels in non-monotonic logics it is crucial that every world (represented by a component of the hypersequent) has access to *all* other worlds which have been constructed so far. This very strongly suggests a *flat* structure, as given by hypersequents.

Structural properties and syntactic completeness. We now investigate the structural properties of our calculi. We first show that weakening and contraction are height-preserving (*hp* for short) admissible, both in their internal and in their external variants. Then, we prove that the cut rule is admissible, which allows us to directly prove the completeness of the calculi with respect to the corresponding axiomatisations. In the proofs we use the following definition of weight of formulas and blocks.

Definition 3. *The weight w of a formula is recursively defined as $w(\perp) = w(\top) = w(p) = 0$, $w(A \rightarrow B) = w(A) + w(B) + 1$, $w(\langle A_1, \dots, A_n \rangle) = \max_i \{w(A_i)\} + 1$, $w(\Box A) = w(A) + 2$.*

Proposition 2. *The following structural rules are hp-admissible in \mathcal{H}_{E^*} :*

$$\begin{array}{cccc} \text{wk}_L \frac{G \mid \Gamma \Rightarrow \Delta}{G \mid A, \Gamma \Rightarrow \Delta} & \text{wk}_R \frac{G \mid \Gamma \Rightarrow \Delta}{G \mid \Gamma \Rightarrow \Delta, A} & \text{wk}_\diamond \frac{G \mid \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} & \text{ctr}_L \frac{G \mid A, A, \Gamma \Rightarrow \Delta}{G \mid A, \Gamma \Rightarrow \Delta} \\ \text{ctr}_R \frac{G \mid \Gamma \Rightarrow \Delta, A, A}{G \mid \Gamma \Rightarrow \Delta, A} & \text{ctr}_\diamond \frac{G \mid \langle \Sigma \rangle, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} & \text{ew} \frac{G}{G \mid \Gamma \Rightarrow \Delta} & \text{ec} \frac{G \mid \Gamma \Rightarrow \Delta \mid \Gamma \Rightarrow \Delta}{G \mid \Gamma \Rightarrow \Delta} \end{array}$$

Note that, since weakening is admissible, invertibility of all rules is immediate.

We now move our attention to the admissibility of the following cut rule

$$\text{cut} \frac{G \mid \Gamma \Rightarrow \Delta, A \quad G \mid A, \Gamma \Rightarrow \Delta}{G \mid \Gamma \Rightarrow \Delta}$$

In order do to this, we prove simultaneously the admissibility of *cut* and of the following rule *sub*, which states that a formula A inside one or more blocks can be replaced by any equivalent set of formulas Σ :

$$\text{sub} \frac{G \mid \Sigma \Rightarrow A \quad \{G \mid A \Rightarrow B\}_{B \in \Sigma} \quad G \mid \langle A^{n_1}, \Pi_1 \rangle, \dots, \langle A^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma^{n_1}, \Pi_1 \rangle, \dots, \langle \Sigma^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta}$$

where A^{n_i} (resp. Σ^{n_i}) is a compact way to denote n_i occurrences of A (resp. Σ). In the monotonic case we need to consider, instead of **sub**, the rule

$$\text{sub}_M \frac{G \mid \Sigma \Rightarrow A \quad G \mid \langle A^{n_1}, \Pi_1 \rangle, \dots, \langle A^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma^{n_1}, \Pi_1 \rangle, \dots, \langle \Sigma^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta}$$

Theorem 2. *If \mathcal{H}_{E^*} contains \square_R , then the rules **cut** and **sub** are admissible in \mathcal{H}_{E^*} , otherwise **cut** and **sub**_M are admissible in \mathcal{H}_{E^*} .*

Sketch of Proof. Let $Cut(c, h)$ mean that all applications of **cut** of height h on a cut formula of weight c are admissible, and $Sub(c)$ mean that all applications of **sub** where A has weight c are admissible. Then the theorem is a consequence of the following claims (for all $\Sigma, \Pi_1, \dots, \Pi_k$): **(A)** $\forall c. Cut(c, 0)$; **(B)** $\forall h. Cut(0, h)$; **(C)** $\forall c. (\forall h. Cut(c, h) \rightarrow Sub(c))$; **(D)** $\forall c. \forall h. ((\forall c' < c. (Sub(c') \wedge \forall h'. Cut(c', h')) \wedge \forall h'' < h. Cut(c, h'')) \rightarrow Cut(c, h))$. Proof in the Appendix. \square

Theorem 3 (Syntactic completeness). *If $i(\Gamma \Rightarrow \Delta)$ is derivable in E^* , then $\Gamma \Rightarrow \Delta$ is derivable in \mathcal{H}_{E^*} .*

Proof. By deriving the axioms, simulating MP using **cut**, and checking that RE is derivable using **ew**, \square_R and \square_L . \square

4 Complexity of proof search

We would like to use the calculus to obtain an optimal decision procedure for the considered logics. As established in [20], logics without axiom C are coNP-complete, whereas logics with C are in PSPACE (although we are not aware of a proof of the lower bound). We accordingly distinguish cases whether axiom C is present or not.

Extensions without C. The decision procedures for the logics **E**, **M**, **EN** and **MN** implement backwards proof search on a polynomially bounded nondeterministic Turing machine with universal choices to handle the branching caused by rules with several premisses, as shown in Alg. 1. To prevent loops, we employ a *local loop checking* strategy, stating that a rule is not applied (bottom-up) to a hypersequent G , if for at least one of its premisses H we have that for every component $\Gamma \Rightarrow \Delta$ in H there is a component $\Sigma \Rightarrow \Pi$ in G with $\text{set}(\Gamma) \subseteq \text{set}(\Sigma)$ and $\text{set}(\Delta) \subseteq \text{set}(\Pi)$. The order of applications of the rules is arbitrary but fixed, and once a rule has been applied the algorithm universally chooses one of its premisses and then recursively checks that this premiss is derivable.

It is easy to see that Alg. 1 is correct and complete. In particular, completeness follows from admissibility of the structural rules of Prop. 2, and the fact that in view of this it suffices to search for *minimal derivations*, i.e., derivations where none of the branches can be shortened.

Algorithm 1: Decision procedure for the derivability problem in \mathcal{H}_{E^*}

Input: a hypersequent G and the code of a logic $\mathbf{L} \in \{\mathbf{E}, \mathbf{M}, \mathbf{EN}, \mathbf{MN}\}$

Output: Is G derivable in $\mathcal{H}_{\mathbf{L}}$?

- 1 **if** there is a component $\Gamma \Rightarrow \Delta$ in G with $\perp \in \Gamma$, or $\top \in \Delta$, or $\Gamma \cap \Delta \neq \emptyset$ **then**
 - 2 | halt and accept;
 - 3 pick the next applicable rule R from $\mathcal{H}_{\mathbf{L}}$, a matching component $\Gamma \Rightarrow \Delta$ and principal formula (and block) from $\Gamma \Rightarrow \Delta$;
 - 4 universally choose a premiss H of this rule application ;
 - 5 check recursively whether H is derivable, output the answer and halt;
 - 6 halt and reject;
-

Theorem 4. *Alg. 1 runs in coNP, whence for the logics without C the calculi provide a complexity-optimal decision procedure.*

Proof. Since the procedure is in the form of a non-deterministic Turing machine with universal choices, it suffices to show that every computation of this machine has polynomial length. Every application of a rule adds either a subformula of its conclusion or a new block to one of the components, or adds a new component. Due to local loop checking it never adds a formula, block or component which is already in the premiss, so it suffices to calculate the maximal size of a hypersequent occurring in proof search for G . Suppose that the size of G is n . Then both the number of components and the number of subformulas of G are bounded by n . Since the local loop check prevents the duplication of formulas, each component contains at most n formulas in the antecedent and n formulas in the succedent. Moreover, since we only consider logics without the axiom C, every newly created block contains exactly one formula. Again, due to the local loop checking condition no block is duplicated, so every component contains at most n blocks. Thus every component has size at most $3n$. The procedure creates new components from a block and a formula of an already existing component, hence there are only n^2 many different components which can be created without violating the local loop checking condition. Thus every hypersequent occurring in the proof contains at most $n + n^2$ many components, each of size at most $3n$, giving a total size and thus running time of $\mathcal{O}(n^3)$. \square

Extensions with C. For the logics with axiom C we would like to use our calculi to obtain PSPACE decision procedures. It can be easily shown that Alg. 1 works properly also for these logics, ensuring in particular termination. However, because of C-rule, hypersequents can be exponentially large, and therefore PSPACE upper bound *cannot* be achieved. In order to obtain PSPACE decision procedures we adopt a different strategy: Instead of the rules in Fig. 1, we consider their *unkleene*'d version, *i.e.*, the ones with all principal formulas and structures deleted from the premisses. For instance $\mathbf{M}\square_{\mathbf{R}}$, $\square_{\mathbf{R}}$ and C are replaced respectively with

$$\frac{G|\Gamma \Rightarrow \Delta|\Sigma \Rightarrow B}{G|\langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B} \quad \frac{G|\Gamma \Rightarrow \Delta|\Sigma \Rightarrow B \quad \{G|\Gamma \Rightarrow \Delta|B \Rightarrow A\}_{A \in \Sigma}}{G|\langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B} \quad \frac{G|\langle \Sigma, \Pi \rangle, \Gamma \Rightarrow \Delta}{G|\langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta}$$

Call the resulting calculus $\mathcal{H}_{E^*}^-$. Backwards proof search is then implemented on an alternating Turing machine by existentially guessing the last applied rule except for **N**, and universally checking that all of its premisses are derivable. To ensure that **N** is applied if it is present in the system, we stipulate that it is applied once to every component of the input, and that if the existentially guessed rule is one of \Box_R or \Box_M , the rule **N** is applied immediately afterwards to each of its premisses. Since no rule application keeps the principal formulas in the premisses, and since the rule **N** if present is applied exactly once to every component, there is no need for any loop checking condition.

The calculi $\mathcal{H}_{E^*}^-$ are sound and complete. Soundness is obvious, since we can add the missing formulas and structures and recover derivations in \mathcal{H}_{E^*} . Completeness can be proved syntactically by a cut elimination argument similar to the one in the previous section, or alternatively by simulating the calculi in [11]. While it is easy to see that the calculi in [11] give a PSPACE upper bound, this is less obvious for $\mathcal{H}_{E^*}^-$ because of the hypersequent structure. Nonetheless we obtain the following result.

Theorem 5. *Backwards proof search in $\mathcal{H}_{E^*}^-$ is in PSPACE.*

Proof. We need to show that every run of the procedure terminates in polynomial time. Assume that the size of the input is n . Let the *weight* of a component in a hypersequent be the sum of the weights of the formulas and blocks occurring in it according to Def. 3, and suppose that the maximal weight of components in the input is w . Then every rule apart from **N** decreases the weight of the component active in its conclusion. Moreover, a new component is only introduced in place of a subformula of the input, hence any hypersequent occurring in the proof search has at most $n+n$ components. The weight of each of these components is at most the maximal weight of a component of the input (plus one in the cases with **N**). Since the rule **N** is applied at most once to each component, it is thus applied at most n times in the total proof search. Thus the runtime in total is $\mathcal{O}(n^2 \cdot w)$, hence polynomial in the size of the input. Thus the procedure runs in alternating polynomial time, and thus in PSPACE. \square

5 Countermodel extraction

We now prove semantic completeness of the calculus, *i.e.*: every valid hypersequent is derivable in the calculus. This amounts to show that a non-provable hypersequent has a countermodel. Countermodels are found in the bi-neighbourhood semantics, as it is more suited for direct countermodels extraction from failed proofs than the standard semantics. The reason is that in order to define a neighbourhood model we need to determine *exactly* the truth sets of formulas: If we want a world w to force $\Box A$, then we have to make sure that $[A]$ belongs to $\mathcal{N}(w)$, thus $[A]$ must be computed. On the contrary, in bi-neighbourhood models it suffices to find a suited pair (α, β) such that $\alpha \subseteq [A]$ and $\beta \subseteq [\neg A]$. As

we shall see, such a pair can be extracted directly from the failed proof even without knowing exactly the extension of $[A]$.

In order to prove semantic completeness we make use of the backwards proof search strategy based on local loop checking already considered in Section 4 (Alg. 1). This strategy amounts to consider the following notion of saturation, stating that a bottom-up application of a rule R is not allowed to a hypersequent G if G already fulfills the corresponding saturation condition (R).

Definition 4 (Saturated hypersequent). *Let $H = \Gamma_1 \Rightarrow \Delta_1 \mid \dots \mid \Gamma_n \Rightarrow \Delta_n$ be a hypersequent occurring in a proof for H' . The saturation conditions associated to each application of a rule of \mathcal{H}_{E^*} are as follows: (init) $\Gamma_i \cap \Delta_i = \emptyset$; (\perp_L) $\perp \notin \Gamma_i$; (\top_R) $\top \notin \Delta_i$; (\rightarrow_L) If $A \rightarrow B \in \Gamma_i$, then $A \in \Delta_i$ or $B \in \Gamma_i$; (\rightarrow_R) If $A \rightarrow B \in \Delta_i$, then $A \in \Gamma_i$ and $B \in \Delta_i$; (\wedge_L) If $A \wedge B \in \Gamma_i$, then $A \in \Gamma_i$ and $B \in \Gamma_i$; (\wedge_R) If $A \wedge B \in \Delta_i$, then $A \in \Delta_i$ or $B \in \Delta_i$; (\Box_L) If $\Box A \in \Gamma_i$, then $\langle A \rangle \in \Gamma_i$; (N) $\langle \top \rangle \in \Gamma_i$; (C) If $\langle \Sigma \rangle, \langle \Pi \rangle \in \Gamma_i$, then there is $\langle \Omega \rangle \in \Gamma_i$ such that $\text{set}(\Sigma, \Pi) = \text{set}(\Omega)$. ($M\Box_R$) If $\langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B$ is in H , then there is $\Sigma, \Gamma' \Rightarrow \Delta', B$ in H . (\Box_R) If $\langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B$ is in H , then there is $\Sigma, \Gamma' \Rightarrow \Delta', B$ in H or there is $B, \Gamma' \Rightarrow \Delta', A$ in H for some $A \in \Sigma$.*

We say that H is saturated with respect to an application of a rule R if it satisfies the corresponding saturation condition (R) for that particular rule application, and that it is saturated with respect to \mathcal{H}_{E^*} if it is saturated with respect to all possible applications of any rule of \mathcal{H}_{E^*} .

Given a saturated hypersequent H we can construct directly a countermodel for H in the bi-neighbourhood semantics in the following way.

Definition 5 (Countermodel construction). *Let H be a saturated hypersequent occurring in a proof for H' . Moreover, let $e : \mathbb{N} \rightarrow H$ be an enumeration of the components of H . Given e , we can write H as $\Gamma_1 \Rightarrow \Delta_1 \mid \dots \mid \Gamma_k \Rightarrow \Delta_k$. Model $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}, \mathcal{V} \rangle$ is defined as follows:*

- $\mathcal{W} = \{n \mid \Gamma_n \Rightarrow \Delta_n \in H\}$.
- $\mathcal{V}(p) = \{n \mid p \in \Gamma_n\}$.
- For all blocks $\langle \Sigma \rangle$ appearing in a component $\Gamma_m \Rightarrow \Delta_m$ of H ,
 $\Sigma^+ = \{n \mid \text{set}(\Sigma) \subseteq \Gamma_n\}$ and $\Sigma^- = \{n \mid \Sigma \cap \Delta_n \neq \emptyset\}$.
 - Non-monotonic case: $\mathcal{N}(n) = \{(\Sigma^+, \Sigma^-) \mid \langle \Sigma \rangle \in \Gamma_n\}$.
 - Monotonic case: $\mathcal{N}(n) = \{(\Sigma^+, \emptyset) \mid \langle \Sigma \rangle \in \Gamma_n\}$.

Lemma 1. *Let \mathcal{M} be defined as in Def. 5. Then for every A , $\langle \Sigma \rangle$ and every $n \in \mathcal{W}$, we have: If $A \in \Gamma_n$, then $n \Vdash A$; if $\langle \Sigma \rangle \in \Gamma_n$, then $n \Vdash \Box \wedge \Sigma$; and if $A \in \Delta_n$, then $n \not\Vdash A$. Moreover, (a) \mathcal{M} is a M -model if \mathcal{H}_{E^*} contains rule $M\Box_R$; (b) \mathcal{M} is a N -model if \mathcal{H}_{E^*} contains rule N ; and (c) \mathcal{M} is a C -model if \mathcal{H}_{E^*} contains rule C .*

Proof. The first claim is proved by mutual induction on A and $\langle \Sigma \rangle$. We only consider the cases of modal formulas, the other are similar and simpler. ($\langle \Sigma \rangle \in \Gamma_n$) In the non-monotonic case, by def $(\Sigma^+, \Sigma^-) \in \mathcal{N}(n)$. We show that $\Sigma^+ \subseteq$

$[\bigwedge \Sigma]$ and $\Sigma^- \subseteq [\neg \bigwedge \Sigma]$, which implies $n \Vdash \Box \bigwedge \Sigma$. If $m \in \Sigma^+$, then $\text{set}(\Sigma) \subseteq \Gamma_m$. By i.h. $m \Vdash A$ for all $A \in \Sigma$, then $m \Vdash \bigwedge \Sigma$. If $m \in \Sigma^-$, then there is $B \in \Sigma \cap \Delta_m$. By i.h. $m \not\Vdash B$, then $m \not\Vdash \bigwedge \Sigma$. In the monotonic case the proof is analogous. ($\Box B \in \Gamma_n$) By saturation of \Box_L , $\langle B \rangle \in \Gamma_n$. Then by i.h. $n \Vdash \Box B$. ($\Box B \in \Delta_n$) In the non-monotonic case, assume $(\Sigma^+, \Sigma^-) \in \mathcal{N}(n)$. Then there is $\langle \Pi \rangle \in \Gamma_n$ s.t. $\Pi^+ = \Sigma^+$ and $\Pi^- = \Sigma^-$. By saturation of \Box_R , there is $m \in \mathcal{W}$ s.t. $\Pi \subseteq \Gamma_m$ and $B \in \Delta_m$, or there is $m \in \mathcal{W}$ s.t. $B \in \Gamma_m$ and $\Pi \cap \Delta_m \neq \emptyset$. In the first case, $m \in \Pi^+$ and by i.h. $m \not\Vdash B$. In the second case, $m \in \Pi^-$ and by i.h. $m \Vdash B$. That is $\Sigma^+ \not\subseteq [B]$ or $\Sigma^- \not\subseteq [\neg B]$. Then $n \not\Vdash \Box B$. The monotonic case is analogous.

The model conditions are proved as follows: (a) By definition of \mathcal{N} in monotonic case. (b) By saturation $\langle \top \rangle \in \Gamma_n$ for all $n \in \mathcal{W}$. Then $(\top^+, \top^-) \in \mathcal{N}(n)$, where by saturation of \top_R , $\top^- = \emptyset$. (c) Assume $(\Sigma^+, \Sigma^-), (\Pi^+, \Pi^-) \in \mathcal{N}(n)$. Then there are $\langle \Lambda \rangle, \langle \Theta \rangle \in \Gamma_n$ s.t. $\Sigma^+ = \Lambda^+$, $\Sigma^- = \Lambda^-$, $\Pi^+ = \Theta^+$ and $\Pi^- = \Theta^-$. By saturation, there is $\langle \Omega \rangle \in \Gamma_n$ s.t. $\text{set}(\Omega) = \text{set}(\Lambda, \Theta)$, thus $(\Omega^+, \Omega^-) \in \mathcal{N}(n)$. We show that (i) $\Omega^+ = \Sigma^+ \cap \Pi^+$ and (ii) $\Omega^- = \Sigma^- \cap \Pi^-$. (i) If $m \in \Omega^+$, then $\text{set}(\Omega) \subseteq \Gamma_m$, then $\text{set}(\Lambda, \Theta) \subseteq \Gamma_m$, then $\text{set}(\Lambda) \subseteq \Gamma_m$ and $\text{set}(\Theta) \subseteq \Gamma_m$, then $m \in \Lambda^+$ and $m \in \Theta^+$. If $m \in \Lambda^+ \cap \Theta^+$, then $m \in \Lambda^+$ and $m \in \Theta^+$, then $\text{set}(\Lambda) \subseteq \Gamma_m$ and $\text{set}(\Theta) \subseteq \Gamma_m$, then $\text{set}(\Lambda, \Theta) \subseteq \Gamma_m$, then $\text{set}(\Omega) \subseteq \Gamma_m$, then $m \in \Omega^+$. (ii) If $m \in \Omega^-$, then $\Omega \cap \Delta_m \neq \emptyset$, then $\Lambda, \Theta \cap \Delta_m \neq \emptyset$, then $\Lambda \cap \Delta_m \neq \emptyset$ or $\Theta \cap \Delta_m \neq \emptyset$, then $m \in \Lambda^-$ or $m \in \Theta^-$. If $m \in \Lambda^- \cup \Theta^-$, then $m \in \Lambda^-$ or $m \in \Theta^-$, then $\Lambda \cap \Delta_m \neq \emptyset$ or $\Theta \cap \Delta_m \neq \emptyset$, then $\Lambda, \Theta \cap \Delta_m \neq \emptyset$, then $\Omega \cap \Delta_m \neq \emptyset$, then $m \in \Omega^-$. \square

Observe that, since all rules are cumulative, \mathcal{M} is a countermodel for the root hypersequent H' . Moreover, since every proof built in accordance to the strategy either provides a derivation of the root hypersequent or contains a saturated hypersequent, this allows us to prove the following theorem.

Theorem 6. *If H is valid in $(M/C/N)$ -models, then it is derivable in $\mathcal{H}_{E(M/C/N)}$.*

As the above construction shows, we can directly extract a bi-neighbourhood countermodel from any failed proof. If we want to obtain a countermodel in the standard semantics we then need to apply the transformations presented in Section 2. In principle, the rough transformation can be embedded into the countermodel construction in order to get immediately a neighbourhood model. However, as illustrated in Section 2 we might obtain a much larger model than needed. On the other hand, there is no obvious way to integrate the finer transformation of Prop. 1 since it rests on the evaluation in an already existing model. An alternative way to obtain countermodels in the neighbourhood semantics is proposed in [11]: It basically consists in forcing the proof search procedure to determine exactly the truth set of each formula. To this aim, whenever a sequent representing a new world is created, the sequent is saturated with respect to all disjunctions $A \vee \neg A$ such that A is a subformula of the root sequent. This solution is equivalent to use *analytic cut* and makes the proof search procedure significantly more complex than the one presented here (moreover it makes use of a more complex data structure than hypersequents).

Below we show some examples of countermodel extraction both in the bi-neighbourhood and in the standard neighbourhood semantics, the latter obtained by Prop. 1.

Example 2 (Proof search for axiom M in \mathcal{H}_E and countermodels).

$$\frac{\frac{\text{closed} \quad \dots \mid p \wedge q \Rightarrow p}{\dots \mid p \wedge q \Rightarrow p} \quad \frac{\text{closed} \quad \dots \mid p \Rightarrow p \wedge q, p \quad \langle p \wedge q \rangle, \Box(p \wedge q) \Rightarrow \Box p \mid p \Rightarrow p \wedge q, q}{\langle p \wedge q \rangle, \Box(p \wedge q) \Rightarrow \Box p \mid p \Rightarrow p \wedge q} \wedge_R \quad \text{saturated hypersequent } H}{\frac{\langle p \wedge q \rangle, \Box(p \wedge q) \Rightarrow \Box p}{\Box(p \wedge q) \Rightarrow \Box p} \Box_L} \Box_R$$

Bi-neighbourhood countermodel. Let us consider the enumeration for the components of \mathcal{B} where $1 \mapsto \langle p \wedge q \rangle, \Box(p \wedge q) \Rightarrow \Box p$, and $2 \mapsto p \Rightarrow p \wedge q, q$. According to the construction in Def. 5, from H we obtain the following countermodel $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}, \mathcal{V} \rangle$: $\mathcal{W} = \{1, 2\}$, $\mathcal{V}(p) = \{2\}$, $\mathcal{V}(q) = \emptyset$, $\mathcal{N}(2) = \emptyset$, and $\mathcal{N}(1) = \{(\emptyset, \{2\})\}$, as $\mathcal{N}(1) = \{(p \wedge q^+, p \wedge q^-)\}$ and $p \wedge q^+ = \emptyset$, $p \wedge q^- = \{2\}$.

Neighbourhood countermodel. Considering $\mathcal{S} = \{\Box(p \wedge q), \Box p, p \wedge q, p, q\}$, by Prop. 1 we obtain from \mathcal{M} the standard model \mathcal{M}' with $\mathcal{N}'(1) = \{\emptyset\}$, as $\mathcal{N}'(1) = \{[p \wedge q]_{\mathcal{M}}\}$ and $[p \wedge q]_{\mathcal{M}} = \emptyset$.

Example 3 (Proof search for axiom K in \mathcal{H}_{EC} and countermodels). By bottom-up proof search for $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ in \mathcal{H}_{EC} we obtain the following branch ending with a saturated hypersequent H (for lack of space we do not show the whole tree).

$$\frac{\frac{\frac{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q \mid q \Rightarrow p \mid p \rightarrow q \Rightarrow q, p}{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q \mid q \Rightarrow p \mid p \rightarrow q \Rightarrow q} \rightarrow_L}{\frac{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q \mid q \Rightarrow p}{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q \mid q \Rightarrow p} \Box_R} \Box_R}{\frac{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q}{\Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle \Rightarrow \Box q} C} \Box_L^2$$

Bi-neighbourhood countermodel. We consider the following enumeration of the components of H : $1 \mapsto \Box(p \rightarrow q), \Box p, \langle p \rightarrow q \rangle, \langle p \rangle, \langle p \rightarrow q, p \rangle \Rightarrow \Box q$; $2 \mapsto q \Rightarrow p$; $3 \mapsto p \rightarrow q \Rightarrow q, p$. From H we obtain $\mathcal{M} = \langle \mathcal{W}, \mathcal{N}, \mathcal{V} \rangle$, where $\mathcal{W} = \{1, 2, 3\}$, $\mathcal{V}(p) = \emptyset$, $\mathcal{V}(q) = \{2\}$, $\mathcal{N}(2) = \mathcal{N}(3) = \emptyset$, and $\mathcal{N}(1) = \{(\emptyset, \{2, 3\}), (\{3\}, \emptyset)\}$, as $\mathcal{N}(1) = \{(p^+, p^-), (p \rightarrow q^+, p \rightarrow q^-), (p, p \rightarrow q^+, p, p \rightarrow q^-)\}$ and $p^+ = \emptyset$, $p^- = \{2, 3\}$, $p \rightarrow q^+ = \{3\}$, $p \rightarrow q^- = \emptyset$, $p, p \rightarrow q^+ = \emptyset$, $p, p \rightarrow q^- = \{2, 3\}$. It is easy to verify that \mathcal{M} is a C-model.

Neighbourhood countermodel. By logical equivalence we can take $\mathcal{S} = \{\Box(p \rightarrow q), \Box p, \Box q, p \rightarrow q, p, q, \Box((p \rightarrow q) \wedge p), \Box(p \wedge q)\}$. We obtain the standard model \mathcal{M}' with $\mathcal{N}'(1) = \{\emptyset, \mathcal{W}\}$.

6 Extensions with axioms T, P, and D

We aim to extend our calculi to systems containing further modal axioms. As a starting point, we consider in this section extensions of non-normal modal logics with axioms T, P, and D:

$$\text{T} \quad \Box A \rightarrow A \quad \text{P} \quad \neg \Box \perp \quad \text{D} \quad \neg(\Box A \wedge \Box \neg A)$$

T is a standard axiom, and P and D are of specific interest in deontic logic: If $\Box A$ is read as ‘ A is obligatory’, then axiom D means that there cannot be two contradicting obligations, whereas axiom P means that there cannot be inconsistent obligations. It is worth noticing that axioms D and P are equivalent in normal modal logics, but are not necessarily equivalent in non-normal ones. The following dependencies hold: $\vdash_{\mathbf{ET}} \text{D}$; $\vdash_{\mathbf{ET}} \text{P}$; $\vdash_{\mathbf{MD}} \text{P}$; $\vdash_{\mathbf{END}} \text{P}$; $\vdash_{\mathbf{ECP}} \text{D}$.

Like the systems of the classical cube, their extensions with axioms T, P, D can be characterised by certain classes of bi-neighbourhood models ([4, 5]). The corresponding conditions are (T) if $(\alpha, \beta) \in \mathcal{N}(w)$, then $w \in \alpha$; (P) if $(\alpha, \beta) \in \mathcal{N}(w)$, then $\alpha \neq \emptyset$; and (D) if $(\alpha, \beta), (\gamma, \delta) \in \mathcal{N}(w)$, then $\alpha \cap \gamma \neq \emptyset$ or $\beta \cap \delta \neq \emptyset$.

Sequent calculi for NNMLs containing axioms T, P, or D have been studied in [9, 10, 15], although in none of them semantic completeness and countermodel extraction are considered. Here we define hypersequent calculi by the rules below.

$$\begin{array}{c} \text{T} \frac{G \mid \Sigma, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \quad \text{P} \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Sigma \Rightarrow}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \quad \text{D}_M \frac{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta \mid \Sigma, \Pi \Rightarrow}{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta} \\ \text{D} \frac{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta \mid \Sigma, \Pi \Rightarrow \quad \{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta \mid \Rightarrow A, B\}_{A \in \Sigma, B \in \Pi}}{G \mid \langle \Sigma \rangle, \langle \Pi \rangle, \Gamma \Rightarrow \Delta} \end{array}$$

The above rules allow us to derive the corresponding axioms as follows:

$$\begin{array}{c} \frac{A, \langle A \rangle, \Box A \Rightarrow A}{\langle A \rangle, \Box A \Rightarrow A} \text{T} \quad \frac{\langle \perp \rangle, \Box \perp \Rightarrow \mid \perp \Rightarrow}{\langle \perp \rangle, \Box \perp \Rightarrow} \text{P} \\ \frac{\langle A \rangle, \langle \neg A \rangle, \Box A, \Box \neg A \Rightarrow \mid A, \neg A \Rightarrow}{\langle A \rangle, \langle \neg A \rangle, \Box A, \Box \neg A \Rightarrow} \text{D} \\ \frac{\langle A \rangle, \langle \neg A \rangle, \Box A, \Box \neg A \Rightarrow \mid \Rightarrow A, \neg A}{\langle A \rangle, \langle \neg A \rangle, \Box A, \Box \neg A \Rightarrow} \text{D} \\ \frac{\langle A \rangle, \langle \neg A \rangle, \Box A, \Box \neg A \Rightarrow}{\Box A, \Box \neg A \Rightarrow} \text{D} \end{array}$$

Cut-free calculi for systems *containing* axioms T or P are obtained just by adding T rule or P rule. For non-monotonic logics containing D (but neither T nor P) it *seems necessary* (at present) to add $\text{ctr}_\langle \rangle$ or the contracted version of D:

$$\text{D}_{\text{aux}} \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Sigma \Rightarrow \quad \{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Rightarrow A\}_{A \in \Sigma}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}$$

In contrast, for monotonic logics containing D we can add rule P (notice that axiom P is derivable in **MD**). As before, one can prove soundness, syntactic completeness, and semantic completeness of hypersequent calculi. Since the calculi are defined modularly, it suffices to extend the proofs in previous sections.

Theorem 7. *If H is derivable in $\mathcal{H}_{E(\mathsf{T}/\mathsf{P}/\mathsf{D})^*}$, then it is valid in the corresponding bi-neighbourhood models.*

Proof. T) Assume $G \mid \Sigma, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta$ valid. Then for all \mathcal{M} , $\mathcal{M} \models G$ or $\mathcal{M} \models \Sigma, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta$. In the second case, for all wolds w of \mathcal{M} , $w \Vdash i(\Sigma, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta)$, which is equivalent to $\bigwedge \Sigma \wedge \Box \bigwedge \Sigma \rightarrow i(\Gamma \Rightarrow \Delta)$. Since \mathcal{M} is a T -model, this is equivalent to $\Box \bigwedge \Sigma \rightarrow i(\Gamma \Rightarrow \Delta)$. Thus $w \Vdash i(\langle \Sigma \rangle, \Gamma \Rightarrow \Delta)$.

P) Assume $G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Sigma \Rightarrow$ valid. If $\mathcal{M} \models G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta$ we are done. Otherwise $\mathcal{M} \models \Sigma \Rightarrow$, that is $[\bigwedge \Sigma] = \emptyset$. Since \mathcal{M} is a P -model, $w \not\Vdash \Box \bigwedge \Sigma$ for all worlds w of \mathcal{M} , which implies $\mathcal{M} \models \langle \Sigma \rangle, \Gamma \Rightarrow \Delta$.

The cases D , D_M and D_{aux} are analogous. \square

It is possible to prove the admissibility of structural rules in Prop. 2. In particular, in case of calculi containing rules for D one may need to use the auxiliary rules D_{aux} and P in order to show admissibility of contraction. As before, this allows us to prove cut elimination, whence syntactic completeness of hypersequent calculi.

Theorem 8. *If $\mathcal{H}_{E(\mathsf{T}/\mathsf{P}/\mathsf{D})^*}$ contains \Box_R , then the rules **cut** and **sub** are admissible in $\mathcal{H}_{E(\mathsf{T}/\mathsf{P}/\mathsf{D})^*}$, otherwise **cut** and **sub**_M are admissible in $\mathcal{H}_{E(\mathsf{T}/\mathsf{P}/\mathsf{D})^*}$.*

Sketch of Proof. By extending the proof of Theorem 2. In particular we need to extend point **(C)** (ii) to the cases where the last rule applied in the derivation of $G \mid \langle A^{n_1}, \Pi_1 \rangle, \dots, \langle A^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta$ is T , P , D , or D_{aux} (resp. T , P , or D_M in monotonic case). \square

Theorem 9 (Syntactic completeness). *If $i(\Gamma \Rightarrow \Delta)$ is derivable in $\mathbf{E}(\mathsf{T}/\mathsf{P}/\mathsf{D})^*$, then $\Gamma \Rightarrow \Delta$ is derivable in $\mathcal{H}_{E(\mathsf{T}/\mathsf{P}/\mathsf{D})^*}$.*

In order to get semantic completeness we need to extend the notion of saturation to the rules for T , P , D . The saturation conditions are the obvious ones, for instance the condition corresponding to T is: If $\langle \Sigma \rangle \in \Gamma_i$, then $\Sigma \subseteq \Gamma_i$. Then, given a saturated hypersequent one can define a countermodel \mathcal{M} by means of the same construction as in Def. 5. We can prove the following lemma.

Lemma 2. *Let H be a saturated hypersequent occurring in a proof for H' , and \mathcal{M} be defined as in Def. 5. Then (a) \mathcal{M} is a T -model if \mathcal{H}_{E^*} contains rule T ; (b) \mathcal{M} is a P -model if \mathcal{H}_{E^*} contains rule P ; and (c) \mathcal{M} is a D -model if \mathcal{H}_{E^*} contains rules D and D_{aux} , or it contains rules D_M and P .*

Proof. (a) Assume $(\Sigma^+, \Sigma^-) \in \mathcal{N}(n)$. Then there is $\langle A \rangle \in \Gamma_n$ s.t. $\Sigma^+ = \Lambda^+$, $\Sigma^- = \Lambda^-$. By saturation of T , $\text{set}(A) \subseteq \Gamma_n$, then by definition $n \in \Lambda^+ = \Sigma^+$.

(b) Assume $(\Sigma^+, \Sigma^-) \in \mathcal{N}(n)$. Then there is $\langle A \rangle \in \Gamma_n$ s.t. $\Sigma^+ = \Lambda^+$, $\Sigma^- = \Lambda^-$. By saturation of P , there is $m \in \mathcal{W}$ such that $\text{set}(A) \subseteq \Gamma_m$, then by definition $m \in \Lambda^+ = \Sigma^+$, that is $\Sigma^+ \neq \emptyset$.

(c) Assume $(\Sigma^+, \Sigma^-), (\Pi^+, \Pi^-) \in \mathcal{N}(n)$. If $(\Sigma^+, \Sigma^-) \neq (\Pi^+, \Pi^-)$, then there are $\langle A \rangle, \langle \Theta \rangle \in \Gamma_n$ s.t. $\Sigma^+ = \Lambda^+$, $\Sigma^- = \Lambda^-$, $\Pi^+ = \Theta^+$ and $\Pi^- = \Theta^-$. If \mathcal{H}_{E^*} is non-monotonic, by saturation of D there is $m \in \mathcal{W}$ such that $\text{set}(A, \Theta) \subseteq$

Γ_m or there is $m \in \mathcal{W}$ such that $A, B \in \Delta_m$ for $A \in \Sigma$ and $B \in \Pi$. In the first case, $\text{set}(\Lambda) \subseteq \Gamma_m$ and $\text{set}(\Theta) \subseteq \Gamma_m$, thus by definition $m \in \Lambda^+ = \Sigma^+$ and $m \in \Theta^+ = \Pi^+$, that is $\Sigma^+ \cap \Pi^+ \neq \emptyset$. In the second case, $m \in \Lambda^- = \Sigma^-$ and $m \in \Theta^- = \Pi^-$, that is $\Sigma^- \cap \Pi^- \neq \emptyset$. If instead \mathcal{H}_{E^*} is monotonic, by saturation of \mathbf{D}_M there is $m \in \mathcal{W}$ such that $\text{set}(\Lambda, \Theta) \subseteq \Gamma_m$. Then $\text{set}(\Lambda) \subseteq \Gamma_m$ and $\text{set}(\Theta) \subseteq \Gamma_m$, thus by definition $m \in \Lambda^+ = \Sigma^+$ and $m \in \Theta^+ = \Pi^+$, that is $\Sigma^+ \cap \Pi^+ \neq \emptyset$. The other possibility is that $(\Sigma^+, \Sigma^-) = (\Pi^+, \Pi^-)$. Then there is $\langle \Lambda \rangle \in \Gamma_n$ s.t. $\Sigma^+ = \Lambda^+$ and $\Sigma^- = \Lambda^-$. In the non-monotonic case, by saturation of \mathbf{D}_{aux} there is $m \in \mathcal{W}$ such that $\text{set}(\Lambda) \subseteq \Gamma_m$ or there is $m \in \mathcal{W}$ such that $A \in \Delta_m$ for $A \in \Sigma$. In the first case, by definition $m \in \Lambda^+ = \Sigma^+$. In the second case, $m \in \Lambda^- = \Sigma^-$. Thus $\Sigma^+ \neq \emptyset$ or $\Sigma^- \neq \emptyset$. In the monotonic case we can consider saturation of \mathbf{P} and conclude that $\Sigma^+ \neq \emptyset$. \square

Since every failed proof returns a saturated hypersequent, this implies semantic completeness of hypersequent calculi

Theorem 10. *If H is valid in $(M/C/N/T/P/D)$ -models, then it is derivable in $\mathcal{H}_{E(M/C/N/T/P/D)}$.*

7 Conclusion

In this paper we have provided hypersequent calculi for the cube of classical Non-Normal Modal logics and some deontic extensions. The Hypersequent formulation is possibly the most adequate, in particular for non-monotone non-normal modal logics, as it ensures good semantic, computational, as well as structural properties. First of all, from a failed proof we can easily extract a countermodel (of polynomial size for logics without C) in the bi-neighbourhood semantics, whence in the standard one. The calculi provide a decision procedure of optimal complexity and enjoy syntactic cut elimination. Finally, they have a natural “almost internal” interpretation, as each component of a hypersequent can be read as a formula of the language. In future research, we intend to extend the calculi to further non-normal modal logics obtained by adding standard modal axioms, possibly including also regular logics which have a non standard relational semantics. Moreover, we intend to use the calculi also for metalogical investigation, *e.g.*, for obtaining proof-theoretic constructive proofs of interpolation complementing the general result in [17] and completing [15]. Finally we wish to study the formal relation with other recent calculi in the literature such as [4, 13] in the form of mutual simulation. We also think of implementing our calculi and comparing them with the theorem prover proposed recently in [6].

Appendix

Theorem 2. *If \mathcal{H}_{E^*} contains \square_R , then the rules **cut** and **sub** are admissible in \mathcal{H}_{E^*} , otherwise **cut** and **sub_M** are admissible in \mathcal{H}_{E^*} .*

Proof. We prove that **cut** and **sub** are admissible in non-monotonic \mathcal{H}_{E^*} ; the proof in the monotonic cases is analogous. Recall that, for an application of **cut**, the *cut formula* is the formula which is deleted by that application, while the *cut height* is the sum of the heights of the derivations of the premisses of **cut**.

The theorem is a consequence of the following claims, where $Cut(c, h)$ means that all applications of **cut** of height h on a cut formula of weight c are admissible, and $Sub(c)$ means that all applications of **sub** where A has weight c are admissible (for all $\Sigma, \Pi_1, \dots, \Pi_k$): **(A)** $\forall c. Cut(c, 0)$. **(B)** $\forall h. Cut(0, h)$. **(C)** $\forall c. (\forall h. Cut(c, h) \rightarrow Sub(c))$. **(D)** $\forall c. \forall h. ((\forall c'. (c' < c. (Sub(c') \wedge \forall h'. Cut(c', h')) \wedge \forall h'' < h. Cut(c, h'')) \rightarrow Cut(c, h))$.

(A) deals with applications of **cut** to initial sequents and is trivial.

(B) If the cut formula has weight 0, then it is \perp , \top , or a propositional variable p . In both situations the proof is by complete induction on h . The basic case $h = 0$ is a particular case of **(A)**. For the inductive step, we distinguish three cases.

(i) The cut formula \perp , \top , or p is not principal in the last rule applied in the derivation of the left premiss. By examining all possible rule applications, we show that the application of **cut** can be replaced by one or more applications of **cut** at a smaller height. For instance, assume that the last rule applied is \square_L .

$$\square_L \frac{G \mid \langle A \rangle, \square A, \Gamma \Rightarrow \Delta, \perp}{G \mid \square A, \Gamma \Rightarrow \Delta, \perp} \quad \frac{G \mid \perp, \square A, \Gamma \Rightarrow \Delta}{G \mid \square A, \Gamma \Rightarrow \Delta} \text{cut}$$

The derivation is transformed as follows, with a hp-application of **wk** and an application of **cut** of smaller height.

$$\frac{G \mid \langle A \rangle, \square A, \Gamma \Rightarrow \Delta, \perp \quad \frac{G \mid \perp, \square A, \Gamma \Rightarrow \Delta}{G \mid \perp, \langle A \rangle, \square A, \Gamma \Rightarrow \Delta} \text{wk}_L}{G \mid \langle A \rangle, \square A, \Gamma \Rightarrow \Delta} \text{cut} \quad \square_L \frac{G \mid \langle A \rangle, \square A, \Gamma \Rightarrow \Delta}{G \mid \square A, \Gamma \Rightarrow \Delta}$$

The situation is similar if the last rule in the derivation of the left premiss is applied to some sequent in G .

(ii) The cut formula \perp , \top , or p is not principal in the last rule applied in the derivation of the right premiss. The case is analogous to (i). As an example, suppose that the last rule applied is $\mathsf{M}\square_R$.

$$\frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B, \perp \quad \frac{G \mid \perp, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B \mid \Sigma \Rightarrow B}{G \mid \perp, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B} \mathsf{M}\square_R}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B} \text{cut}$$

The derivation is converted into

$$\text{ew} \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B, \perp}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B, \perp \mid \Sigma \Rightarrow B} \quad \frac{G \mid \perp, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B \mid \Sigma \Rightarrow B}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B \mid \Sigma \Rightarrow B} \text{cut} \quad \mathsf{M}\square_R \frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B \mid \Sigma \Rightarrow B}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \square B}$$

where cut is applied at a smaller height.

(iii) The cut formula \perp , \top , or p is principal in the last rule applied in the derivation of both premisses. Then the cut formula is p , as \perp (resp. \top) is never principal on the right-hand side (resp. left-hand side) of the conclusion of any rule application. This means that both premisses are derived by *init*, which implies $h = 0$. Then we are back to case **(A)**.

(C) Assume $\forall h \text{Cut}(c, h)$. The proof is by induction on the height m of the derivation of $G \mid \langle A^{n_1}, \Pi_1 \rangle, \dots, \langle A^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta$. Here we only consider the case where $m > 0$ and the last rule applied in the derivation is \square_R , with one block among $\langle A, \Pi_1 \rangle, \dots, \langle A, \Pi_k \rangle$ principal in the rule application:

$$\frac{\textcircled{1} \quad \frac{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid D \Rightarrow C \}_{C \in \Pi_i}}{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid A^{n_i}, \Pi_i \Rightarrow D} \quad \frac{\{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid D \Rightarrow A\}_1^{n_i}}{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D} \quad \vdots}{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D} \square_R$$

The derivation is converted as follows. First we derive:

$$\frac{\frac{G \mid \Sigma \Rightarrow A}{G \mid \Sigma \Rightarrow A \mid A^{n_i}, \Pi_i \Rightarrow D} \text{ew} \quad \left\{ \frac{G \mid A \Rightarrow B}{G \mid A \Rightarrow B \mid A^{n_i}, \Pi_i \Rightarrow D} \text{ew} \right\}_{B \in \Sigma} \quad \textcircled{1}}{G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid A^{n_i}, \Pi_i \Rightarrow D} \text{sub}$$

Moreover, by applying *ew* to $G \mid \Sigma \Rightarrow A$ we obtain $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid \Sigma \Rightarrow A$. By auxiliary applications of *wk* we can cut A and get $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid \Sigma, A^{n_i-1}, \Pi_i \Rightarrow D$. Then with further applications of cut (each time with auxiliary applications of *wk*) we obtain $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid \Sigma^{n_i}, \Pi_i \Rightarrow D$. By doing the same with the other premisses of \square_R in the initial derivation we obtain also $\{G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid D \Rightarrow B\}_{B \in \Sigma} (1..n_1)$ and $\{G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D \mid D \Rightarrow C\}_{C \in \Pi_i}$. Then by \square_R we derive the conclusion of *sub* $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma' \Rightarrow \Delta', \square D$.

(D) Assume $\forall c' < c. (\text{Sub}(c') \wedge \forall h'. \text{Cut}(c', h'))$ and $\forall h'' < h. \text{Cut}(c, h'')$. We show that all applications of cut of height h on a cut formula of weight c can be replaced by different applications of cut, either of smaller height or on a cut formula of smaller weight. We can assume $c, h > 0$ as the cases $c = 0$ and $h = 0$ have been considered already in **(B)** and **(A)**. We distinguish two cases.

(i) The cut formula is not principal in the last rule application in the derivation of at least one of the two premisses of cut. This case is analogous to (i) or (ii) in **(B)**.

(ii) The cut formula is principal in the last rule application in the derivation of both premisses. Then the cut formula is either $B \rightarrow C$, or $B \wedge C$, or $\square B$.

— If the cut formula is $B \rightarrow C$ we have

$$\rightarrow_R \frac{\frac{\textcircled{1} G \mid B, \Gamma \Rightarrow \Delta, B \rightarrow C, C}{\textcircled{2} G \mid \Gamma \Rightarrow \Delta, B \rightarrow C} \quad \frac{\textcircled{3} G \mid B \rightarrow C, \Gamma \Rightarrow \Delta, B}{\textcircled{5} G \mid B \rightarrow C, \Gamma \Rightarrow \Delta} \quad \frac{\textcircled{4} G \mid C, B \rightarrow C, \Gamma \Rightarrow \Delta}{\textcircled{5} G \mid B \rightarrow C, \Gamma \Rightarrow \Delta}}{G \mid \Gamma \Rightarrow \Delta} \text{cut} \rightarrow_L$$

The derivation is converted into the following one:

$$\frac{\frac{\text{wk}_R \frac{\text{cut} \frac{\text{wk}_R \frac{G \mid \Gamma \Rightarrow \Delta, B \rightarrow C, B}{\text{cut}}}{G \mid \Gamma \Rightarrow \Delta, B, C}}{G \mid \Gamma \Rightarrow \Delta, C}}{G \mid \Gamma \Rightarrow \Delta} \quad \frac{\text{wk}_L \frac{\text{cut} \frac{\text{wk}_R \frac{G \mid B, B \rightarrow C, \Gamma \Rightarrow \Delta}{\text{cut}}}{G \mid B, B \rightarrow C, \Gamma \Rightarrow \Delta, C}}{G \mid B, \Gamma \Rightarrow \Delta, C}}{G \mid \Gamma \Rightarrow \Delta} \quad \frac{\text{wk}_L \frac{\text{cut} \frac{\text{wk}_L \frac{G \mid C, \Gamma \Rightarrow \Delta, B \rightarrow C}{\text{cut}}}{G \mid C, \Gamma \Rightarrow \Delta}}{G \mid \Gamma \Rightarrow \Delta}}{G \mid \Gamma \Rightarrow \Delta} \text{cut}}{G \mid \Gamma \Rightarrow \Delta} \text{cut}$$

— If the cut formula is $B \wedge C$ the situation is similar.

— If the cut formula is $\Box B$ we have

$$\textcircled{1} \ G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid \Sigma \Rightarrow B$$

$$\square_R \frac{\vdots \quad \frac{\textcircled{2} \ G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B \mid B \Rightarrow C \}_{C \in \Sigma}}{\textcircled{3} \ G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B} \quad \frac{\textcircled{4} \ G \mid \langle B \rangle, \Box B, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}{\textcircled{5} \ G \mid \Box B, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{cut}$$

The derivation is converted as follows, with several applications of cut of smaller height.

$$\frac{\frac{\textcircled{3} \ G \mid \langle B \rangle, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta, \Box B}{G \mid \langle B \rangle, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{wk}_{\langle \rangle} \quad \textcircled{4} \ G \mid \langle B \rangle, \Box B, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}{\textcircled{*} \ G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \langle B \rangle, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{cut}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{ew}$$

$$\frac{\textcircled{1} \ \frac{\textcircled{5} \ G \mid \Box B, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Sigma \Rightarrow B}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \Sigma \Rightarrow B} \text{ew} \quad \left(\frac{\textcircled{2} \ G \mid \Box B, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid B \Rightarrow C}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid B \Rightarrow C} \text{ew} \right)_{C \in \Sigma} \quad \textcircled{*} \ \text{sub}}{\frac{\frac{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \langle \Sigma \rangle, \langle \Sigma \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{ctr}_{\langle \rangle}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{ec}}{G \mid \langle \Sigma \rangle, \Gamma \Rightarrow \Delta} \text{sub}} \quad \square$$

Theorem 8. *If $\mathcal{H}_{E(\mathcal{T}/\mathcal{P}/\mathcal{D})^*}$ contains \square_R , then the rules cut and sub are admissible in $\mathcal{H}_{E(\mathcal{T}/\mathcal{P}/\mathcal{D})^*}$, otherwise cut and sub_M are admissible in $\mathcal{H}_{E(\mathcal{T}/\mathcal{P}/\mathcal{D})^*}$.*

Proof. We extend point (C) (ii) in the proof of Theorem 2 to the cases where the last rule applied in the derivation of $G \mid \langle A^{n_1}, \Pi_1 \rangle, \dots, \langle A^{n_k}, \Pi_k \rangle, \Gamma \Rightarrow \Delta$ is T, P, D, or D_{aux} (resp. T, P, or D_M in monotonic case). We consider as examples the following two cases.

— The last rule is T:

$$\frac{G \mid A^{n_i}, \Pi_i, \langle A^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta}{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta} \text{T}$$

By applying the inductive hypothesis to the premiss we obtain $G \mid A^{n_i}, \Pi_i, \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta$. Then, from this and $G \mid \Sigma \Rightarrow A$, by several applications of cut (each time with auxiliary applications of wk) we obtain $G \mid \Sigma^{n_i}, \Pi_i, \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta$. Finally, by T we derive $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta$.

— The last rule is P:

$$\frac{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta \mid A^{n_i}, \Pi_i \Rightarrow}{G \mid \langle A^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta} \text{P}$$

By applying the inductive hypothesis to the premiss (after auxiliary applications of *ew* to the other premisses of *sub*) we obtain $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta \mid A^{n_i}, \Pi_i \Rightarrow$. Then, from this and $G \mid \Sigma \Rightarrow A$, by several applications of *cut* (each time with auxiliary applications of *wk*) we obtain $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta \mid \Sigma^{n_i}, \Pi_i \Rightarrow$. Finally, by *P* we derive $G \mid \langle \Sigma^{n_i}, \Pi_i \rangle, \Gamma \Rightarrow \Delta$. \square

References

1. Askounis, D., Koutras, C.D., Zikos, Y.: Knowledge means ‘all’, belief means ‘most’. *Journal of Applied Non-Classical Logics* **26**(3), 173–192 (2016)
2. Avron, A.: The method of hypersequents in the proof theory of propositional non-classical logics. In: *Logic: From Foundations to Applications*. Clarendon P. (1996)
3. Chellas, B.F.: *Modal Logic*. Cambridge University Press (1980)
4. Dalmonte, T., Olivetti, N., Negri, S.: Non-normal modal logics: Bi-neighbourhood semantics and its labelled calculi. In: *Proceedings of AiML* (2018)
5. Dalmonte, T., Olivetti, N., Negri, S.: Bi-neighbourhood semantics and labelled sequent calculi for non-normal modal logics. Part 1 (draft) (2019), <http://www.lsis.org/olivetti/TR2019/TR-BINS.pdf>
6. Dalmonte, T., Olivetti, N., Negri, S., Pozzato, G.L.: Pronom: proof-search and countermodel generation for non-normal modal logics. In: *Proceedings of AIIA* (to appear) (2019)
7. Gilbert, D., Maffezoli, P.: Modular sequent calculi for classical modal logics. *Studia Logica* **103**(1), 175–217 (2015)
8. Goble, L.: Prima facie norms, normative conflicts, and dilemmas. *Handbook of deontic logic and normative systems* **1**, 241–352 (2013)
9. Indrzejczak, A.: Sequent calculi for monotonic modal logics. *Bulletin of the Section of logic* **34**(3), 151–164 (2005)
10. Indrzejczak, A.: Admissibility of cut in congruent modal logics. *Logic and Logical Philosophy* **21**, 189–203 (2011)
11. Lavendhomme, R., Lucas, T.: Sequent calculi and decision procedures for weak modal systems. *Studia Logica* **65**, 121–145 (2000)
12. Lellmann, B.: Combining monotone and normal modal logic in nested sequents - with countermodels. In: *TABLEAUX*. pp. 203–220 (2019)
13. Lellmann, B., Pimentel, E.: Modularisation of sequent calculi for normal and non-normal modalities. *ACM Trans. Comput. Logic* **20**(2), 7:1–7:46 (2019)
14. Negri, S.: Proof theory for non-normal modal logics: The neighbourhood formalism and basic results. *IfCoLog J. Log. Appl* **4**(4), 1241–1286 (2017)
15. Orlandelli, E.: Sequent calculi and interpolation for non-normal logics. arXiv preprint arXiv:1903.11342 (2019)
16. Pacuit, E.: *Neighborhood semantics for modal logic*. Springer (2017)
17. Pattinson, D.: The logic of exact covers: Completeness and uniform interpolation. In: *LICS 2013* (2013)
18. Pauly, M.: A modal logic for coalitional power in games. *J. Log. Comput.* **12**(1), 149–166 (2002)
19. Vardi, M.Y.: On epistemic logic and logical omniscience. In: *Theoretical aspects of reasoning about knowledge*. pp. 293–305. Elsevier (1986)
20. Vardi, M.Y.: On the complexity of epistemic reasoning. *Proc. 4th IEEE Symp. on Logic in Computer Science* pp. 243–252 (1989)