Valentini's cut-elimination for provability logic resolved

Rajeev Goré and Revantha Ramanayake

ABSTRACT. In 1983, Valentini presented a syntactic proof of cut-elimination for a sequent calculus GLS_V for the provability logic GL where we have added the subscript V for "Valentini". The sequents in GLS_V were built from sets, as opposed to multisets, thus avoiding an explicit contraction rule. From a syntactic point of view, it is more satisfying and formal to explicitly identify the applications of the contraction rule that are 'hidden' in these set-based proofs of cut-elimination. There is often an underlying assumption that the move to a proof of cut-elimination for sequents built from multisets is easy. Recently, however, it has been claimed that Valentini's arguments to eliminate cut do not terminate when applied to a multiset formulation of GLS_V with an explicit rule of contraction. The claim has led to much confusion and various authors have sought new proofs of cut-elimination for GL in a multiset setting.

Here we refute this claim by placing Valentini's arguments in a formal setting and proving cut-elimination for sequents built from multisets. The formal setting is particularly important for sequents built from multisets, in order to accurately account for the interplay between the weakening and contraction rules. Furthermore, Valentini's original proof relies on a novel induction parameter called "width" which is computed 'globally'. It is difficult to verify the correctness of his induction argument based on "width". In our formulation however, verification of the induction argument is straightforward. Finally, the multiset setting also introduces a new complication in the the case of contractions above cut when the cut-formula is boxed. We deal with this using a new transformation based on Valentini's original arguments.

Finally, we show that the algorithm purporting to show the non-termination of Valentini's arguments is not a faithful representation of the original arguments, but is instead a transformation already known to be insufficient.

Keywords: cut elimination, provability logic, Gödel-Löb logic

1 Introduction

The provability logic GL is obtained by adding Löb's axiom $\Box(\Box A \supset A) \supset \Box A$ to the standard Hilbert calculus for propositional normal modal logic K [11]. Interpreting the modal operator $\Box A$ as the provability predicate "A is provable in Peano arithmetic", it can be shown that GL is complete with respect to the formal provability interpretation in Peano arithmetic (see [14]). For an introduction to provability logic see [13].

In 1981, Leivant [5] proposed a syntactic proof of cut-elimination for a sequent calculus for GL. Valentini [16] soon described a counter-example to this proof, proposing a more complicated proof for the sequent calculus GLS_V for GL. The calculus GLS_V is a sequent calculus for classical propositional logic together with a single modal rule GLR. Valentini's proof appears to be the first proof of cut-elimination for a sequent calculus for GL and relies on a complicated transformation justified by a Gentzen-style induction on the degree of the cut-formula and the cut-height, as well as a new induction parameter — the width of a cut-formula. Roughly speaking, the width of a cut-instance is the number of GLR rule instances above that cut which contain a parametric ancestor of the cut-formula in their conclusion. However, Valentini's proof is very brief, informal and difficult to check. For example, he only considers a cut-instance where the cut-formula is left and right principal by the GLR rule (the Sambin Normal Form), noting that "the presence of the new parameter [width] does not affect the [remainder of the standard cut-elimination proof]" [16]. While it is true that the standard transformations appropriately reduce the degree and/or cut-height, he fails to observe that these transformations can sometimes increase the width of lower cuts, casting doubt on the validity of the induction. A careful study of the proposed transformation is required to confirm that the proof is not affected (see Remark 21).

Several other solutions for cut-elimination have been proposed. Borga [2] presented one solution, while Sasaki [12] described a proof for a sequent calculus very similar to GLS_V, relying on cut-elimination for $K4$. Note that only Leivant and Valentini used traditional Gentzen-style methods involving an induction over the degree of the cut-formula and the cut-height.

All four authors used sequents $X \Rightarrow Y$ where X and Y are *sets*, so these calculi did not require a rule of contraction as there is no notion of a set containing an element multiple times (unlike a multiset where the number of occurrences is important). Thus the following instance of the $L\wedge$ rule is legal in GLS_V even though it 'hides' a contraction on $P \wedge Q$:

$$\frac{P \wedge Q, P \Rightarrow R}{P \wedge Q \Rightarrow R}\; L\wedge$$

From a syntactic viewpoint, it is more satisfying and formal to explicitly identify the contractions that are 'hidden' in these set-based proofs of cut-elimination. The appropriate formalisation to understand the reliance on contraction is to use multisets. But then the contraction rules often pose new problems that require attention. For example, Gentzen [4] in his original proof of the *Hauptsatz* for the classical sequent calculus LK, introduced a 'multicut' rule to deal with a complication in the case of contractions above cut. Nevertheless, even that proof is not purely syntactic in the following sense: since multicut is not a rule of the original calculus, the proof has to detour via a conservative extension. Proofs of purely syntactic cut-elimination for LK have subsequently appeared in the literature (see [10],[3],[1] for example). In the case of GL, it turns out that additional complications also arise when formulating Valentini's arguments in a multiset setting, for example, due to the interplay between weakening and contraction rules (see Remark 16).

Thus the move to a proof of cut-elimination for sequents built from multisets is not straightforward. Moen [7] attempted to lift Valentini's set-based arguments to obtain a proof for sequents built from multisets before concluding that this was not possible. Specifically, he presents a concrete derivation ϵ containing cut, and claims that a multiset formulation of Valentini's argument does not terminate when applied to ϵ. Not surprisingly, this claim has ignited the search for new proofs of purely syntactic cut-elimination in a Gentzen-style multiset setting for GL.

In response, Negri [8] and Mints [6] proposed two different solutions. Negri uses a non-standard multiset sequent calculus in which sequents are built from multisets of labelled formulae of the form $x : A$, where A is a traditional formula and x is an explicit name for a Kripke world. She shows that contraction is height-preserving admissible in this calculus and uses this to handle contractions above cut in her cut-elimination argument. In our view, the use of semantic information in the calculus deviates from a purely proof theoretic approach. Mints [6] solves the problem using a sequent calculus similar to the multiset-formulation of GLS_V, but does not state how to handle contractions above cut.

So there are two issues to consider:

1. formalise "width" more precisely to clarify Valentini's original definition, and check whether it is a suitable induction measure, and
2. determine whether Valentini's arguments can be used to obtain a purely syntactic proof of cut-elimination in a *multiset* setting.

Our contribution is as follows: we have successfully translated Valentini's set-based arguments for cut-elimination to a sequent calculus built from multisets. To this end, we have formalised the notion of parametric ancestor and width to correspond intuitively with Valentini's original definition. With this formalisation we show that Valentini's arguments can be applied in the multiset setting, noting that although certain transformations may increase the width of lower cuts, this does not affect the proof. In the case where the last rule in either premise derivation of the cut-rule is a contraction on the cut-formula, we avoid the multicut rule by using von Plato's arguments [10] when the cut-formula is not boxed, and a new argument for the case when the cut-formula is boxed. Thus we obtain a purely syntactic proof of cut-elimination in a multiset setting. We also believe that we have identified a mistake in Moen's claim that Valentini's arguments (in a multiset setting) do not terminate. It appears that Moen has not used a faithful representation of Valentini's arguments for the inductive case, but instead a transformation he titles Val-II(core) that is already known to be insufficient [11]. We discuss this further in Section 5. Of course, the incorrectness of Moen's claim does not imply the correctness of Valentini's arguments in a multiset setting. Indeed the whole point is that complications do arise in the multiset setting, and that these have to be dealt with carefully.

Initial sequents: $A \Rightarrow A$ for each formula A

Logical rules:

$$\frac{X \Rightarrow Y, A}{X, \neg A \Rightarrow Y} L\neg \qquad \frac{A, X \Rightarrow Y}{X \Rightarrow Y, \neg A} R\neg$$

$$\frac{A_i, X \Rightarrow Y}{A_1 \wedge A_2, X \Rightarrow Y} L\wedge \qquad \frac{X \Rightarrow Y, A_1 \qquad X \Rightarrow Y, A_2}{X \Rightarrow Y, A_1 \wedge A_2} R\wedge$$

$$\frac{A_1, X \Rightarrow Y \qquad A_2, X \Rightarrow Y}{A_1 \vee A_2, X \Rightarrow Y} L\vee \qquad \frac{X \Rightarrow Y, A_i}{X \Rightarrow Y, A_1 \vee A_2} R\vee$$

$$\frac{X \Rightarrow Y, A \qquad B, X \Rightarrow Y}{A \supset B, X \Rightarrow Y} L\supset \qquad \frac{A, X \Rightarrow Y, B}{X \Rightarrow Y, A \supset B} R\supset$$

Modal rule:

$$\frac{\Box X, X, \Box B \Rightarrow B}{\Box X \Rightarrow \Box B} GLR$$

Structural rules:

$$\frac{X \Rightarrow Y}{A, X \Rightarrow Y} LW \qquad \frac{X \Rightarrow Y}{X \Rightarrow Y, A} RW$$

$$\frac{A, A, X \Rightarrow Y}{A, X \Rightarrow Y} LC \qquad \frac{X \Rightarrow Y, A, A}{X \Rightarrow Y, A} RC$$

Cut-rule:

$$\frac{X \Rightarrow Y, D \qquad D, U \Rightarrow W}{X, U \Rightarrow Y, W} cut$$

Table 1. The sequent calculus *GLS*

Finally, we remind the reader that it is trivial to show that the cut-rule is redundant for both set and multiset sequent calculus formulations for *GL* by proving that the calculus without the cut-rule is sound and complete for the Kripke semantics of *GL*. However, our purpose here is to resolve the claim about the failure of *syntactic* cut-elimination based on Valentini's arguments for a sequent calculus built with multisets.

2 Preliminaries

We use the letters $P, Q, \ldots$ to denote propositional variables. Formulae are defined in the usual way in terms of propositional variables, the logical constant $\bot$ and the logical connectives $\wedge$ (conjunction), $\vee$ (disjunction), $\supset$ (implication) and $\Box$ (necessity, or in this context, provability). Formulae are denoted by $A, B, \ldots$. Multisets of formulae are denoted by X, Y, U, V, W, G and also as a list of comma-separated formula enclosed in "$\langle$" and "$\rangle$". A formula A is said to be *boxed* if it is of the form $\Box B$ for some formula B and is *not boxed* otherwise. The relation '$\equiv$' is used to denote syntactic equality. Let X be the multiset $\langle A_1, \ldots, A_n \rangle$. Then we define the multiset $\Box X$ to be $\langle \Box A_1, \ldots, \Box A_n \rangle$. Furthermore $B \in X$ iff $B \equiv A_i$ for some $1 \leq i \leq n$. The negation of $B \in X$ is denoted by $B \notin X$. The notation $(A)^m$ or A^m denotes m comma-separated occurrence of A.

A *sequent* is a tuple (X, Y) of multisets X and Y of formulae and is written $X \Rightarrow Y$. We sometimes use $\mathcal{S}$ or $\mathcal{S}'$ to denote a sequent. The multiset X (Y) is called the antecedent (succedent). The multiset sequent calculus we use here is called GLS (Table 1). For the logical and structural rules in GLS, the multisets X and Y are called the *context*. In the conclusion of each of these rules, the formula occurrence not in the context is called the *principal formula*. This follows standard practice (see [15]). For the GLR rule, each formula in $\Box X, X, \Box B, B$ is called a principal formula. The $\Box B$ in the succedent of the conclusion of the GLR rule is called the *diagonal formula* (and is of course boxed). In the cut-rule, the formula D is the *cut-formula*. A rule with one premise (two premises) is called a unary (binary) rule.

A binary rule where the context in both premises is required to be identical is called an *additive* binary rule (eg: $L\vee, R\wedge$). A binary rule where the context in each premise need not be identical is called a *multiplicative* binary rule (eg: *cut*). The term context-sharing (context-independent) is also used to refer to an additive (multiplicative) rule (see [15]).

Note, we have deleted the initial sequent $\bot \Rightarrow \bot$ and the $\bot$-rule that appears in GLS_V. As [13] observes, it is not necessary to include $\bot$ although its presence can be convenient from a semantic viewpoint. As our concerns here are proof theoretic we shall not require it. We have also replaced the multiplicative $L\supset$ in GLS_V with an additive version. As all the other logical rules in GLS are additive, it seems appropriate to use an additive $L\supset$. We observe that the definitions and proofs in this paper apply, with slight amendment, to a sequent calculus built from multisets that is obtained directly from GLS_V.

A *derivation* (in GLS) is defined recursively with reference to Table 1 as:

(i) an initial sequent $A \Rightarrow A$ for any formula A is a derivation, and

(ii) an application of a logical, modal, structural or cut-rule to derivations concluding its premise(s) is a derivation.

This is the standard definition. Viewing a derivation as a tree, we call the root of the tree the *end-sequent* of the derivation. If there is a derivation with end-sequent $X \Rightarrow Y$ we say that $X \Rightarrow Y$ is *derivable* in GLS. Let $\bigwedge X$ ($\bigvee Y$) denote the conjunction (disjunction) of all formula occurrences in X (Y). Interpreting the sequent $X \Rightarrow Y$ as the formula $\bigwedge X \supset \bigvee Y$, from [11] we see that derivability in GLS is sound and complete wrt GL.

We write $\{\pi\}_1^r/^\rho X \Rightarrow Y$ to denote the derivation

$$\frac{\pi_1 \qquad \ldots \qquad \pi_r}{X \Rightarrow Y}\ \rho$$

where ρ is an instance of a rule with r premises. We refer to $\pi_1, \ldots, \pi_r$ as the *premise derivations* of ρ. If ρ is unary (binary) then $r = 1$ ($r = 2$). Rather than $\{\pi\}_1^1$ and $\{\pi\}_1^2$, we write, respectively, "π_1" and "$\pi_1 \quad \pi_2$".

Let ρ be some rule-occurrence in a derivation τ. Then $\rho(A)$ indicates that the principal formula is A; $\rho^*(X)$ denotes some number (≥ 0) of applications of ρ that make each formula occurrence (including multiple formula occurrences) in the multiset X a principal formula. To identify a rule-occurrence in the text we occasionally use subscripts, eg: GLR_1, cut_0.

A derivation τ is *cut-free* if τ contains no instances of the cut-rule. A cut-instance is said to be *topmost* if its premise derivations are cut-free.

DEFINITION 1 (n-ary GLR rule). Given a derivation τ, an instance ρ of the GLR rule appearing in τ is n-ary if there are exactly $n-1$ GLR rule instances on the path between ρ and the end-sequent of τ.

Let $GLR(n, \tau)$ denote the number of n-ary GLR rules in τ. Next we define the height, cut-height, and degree of a formula in the standard manner.

DEFINITION 2 (height, cut-height, degree). The *height* $s(\tau)$ of a derivation τ is the greatest number of successive applications of rules in it plus one. The *cut-height* h of an instance of the cut-rule with premise derivations τ_1 and τ_2 is $s(\tau_1) + s(\tau_2)$. The *degree* $deg(A)$ of a formula A is defined as the number of symbol occurrences in A from $\{\Box, \neg, \wedge, \vee, \supset\}$ plus one.

3 Generalising the notion of derivation

To formalise the notion of width we need a more general structure than a derivation. The structure we have in mind can be obtained from a derivation τ by deleting a proper subderivation τ' in τ. We call this structure a *stub-derivation*. To emphasise the point of deletion we use the annotation stub.

Formally a stub-derivation (in GLS) is defined recursively with reference to Table 1 as follows:

(i) an initial sequent $A \Rightarrow A$ for any formula A is a stub-derivation, and

(ii) for any sequent $\mathcal{S}$ and stub-derivation π, each of

(a) $\mathsf{stub}/\mathcal{S}$ (b) $\mathsf{stub} \quad \pi/\mathcal{S}$ (c) $\pi \quad \mathsf{stub}/\mathcal{S}$

is a stub-derivation, and

(iii) an application of a logical, modal, structural or cut-rule to stub-derivations concluding its premise(s) is a stub-derivation.

Viewing a stub-derivation τ as a tree, we call the root of the tree the *end-sequent* of the stub-derivation (denoted $ES(\tau)$). The leaves of the tree are called the *top-sequents*. Clearly a derivation is a stub-derivation in which every top-sequent is an initial sequent. Thus a stub-derivation generalises the notion of a derivation.

We use the term 'stub-instance' to refer to an occurrence of either $\mathsf{stub}/\mathcal{S}$ or $\mathsf{stub} \quad \pi/\mathcal{S}$ or $\pi \quad \mathsf{stub}/\mathcal{S}$. An *sstub-derivation* (read: single stub-derivation) is a stub-derivation containing exactly one stub-instance. We write $d[\mathsf{stub}]$ instead of d, to remind the reader that the structure contains exactly one stub-instance.

Let d' be a derivation with end-sequent $\mathcal{S}'$, let $d[\mathsf{stub}]$ be an sstub-derivation with an occurrence of one of the following:

$$\mathsf{stub}/\mathcal{S} \qquad\qquad \mathsf{stub} \quad \pi/\mathcal{S} \qquad\qquad \pi \quad \mathsf{stub}/\mathcal{S}$$

and suppose that

$$\mathcal{S}'/^{\rho}\mathcal{S} \qquad\qquad \mathcal{S}' \quad ES(\pi)/\mathcal{S} \qquad\qquad ES(\pi) \quad \mathcal{S}'/\mathcal{S}$$

respectively is a legal instance of some logical or structural rule ρ. We say that $d[\mathsf{stub}]$ and d' are *compatible* and write $d[\mathsf{stub}] \mapsfrom d'$ to denote

$$\frac{d'}{\mathcal{S}}\,\rho \qquad\qquad \frac{d' \qquad \pi}{\mathcal{S}}\,\rho \qquad\qquad \frac{\pi \qquad d'}{\mathcal{S}}\,\rho$$

respectively, obtained by "attaching" the tree d' to the tree $d[\mathsf{stub}]$ at the node stub under rule ρ. We refer to ρ as a *binding rule* for $d[\mathsf{stub}]$ and d'.

By permitting formula occurrences in a (stub-)derivation to contain $*$ or $\circ$ decorations, we define an *annotated (stub-)derivation.* The forgetful map $\lfloor\cdot\rfloor$ maps an annotated stub-derivation to the stub-derivation obtained by erasing all $*$ and $\circ$ decorations. Clearly $\lfloor\cdot\rfloor$ maps an annotated derivation to a derivation. A *transformed (stub-)derivation* τ' is a (stub-)derivation that is obtained from some existing (stub-)derivation τ by syntactic transformation. We write $A^{\circ n}$ or A^{*n} to mean n occurrences of the formula A° or A^{*} respectively.

Formally a stub-derivation and an annotated stub-derivation are different structures. Because these structures are very similar, for economy of space we will introduce definitions and prove results for stub-derivations alone and note, whenever applicable, that the definitions and results can be extended to annotated stub-derivations.

EXAMPLE 3. Let us denote the sstub-derivation at below left by $d[\mathsf{stub}]$ and the derivation at below right by d':

$$\frac{\dfrac{\mathsf{stub}}{B \Rightarrow A \supset B} \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B}\,L\vee \qquad\qquad \frac{B \Rightarrow B}{A, B \Rightarrow B}\,LW$$

Observe that $d[\mathsf{stub}]$ has a stub-instance of type $\mathsf{stub}/\mathcal{S}$, with $\mathcal{S} \equiv B \Rightarrow A \supset B$, and d' has endsequent $\mathcal{S}' \equiv A, B \Rightarrow B$. Because $\mathcal{S}'/\mathcal{S}$ is an instance of $R\supset$, $d[\mathsf{stub}]$ and d' are compatible. The derivation $d[\mathsf{stub}] \mapsfrom d'$ is:

$$\frac{\dfrac{\dfrac{B \Rightarrow B}{A, B \Rightarrow B}\,LW}{B \Rightarrow A \supset B}\,R\supset \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B}\,L\vee$$

and the binding rule is $R\supset$.

EXAMPLE 4. Let us denote the sstub-derivation at below left by $d[\mathsf{stub}]$ and the derivation at below right by d':

$$\dfrac{\mathsf{stub} \qquad A \supset B \Rightarrow A \supset B}{B \vee (A \supset B) \Rightarrow A \supset B} \qquad\qquad \dfrac{\dfrac{B \Rightarrow B}{A, B \Rightarrow B}\ LW}{B \Rightarrow A \supset B}\ R\supset$$

Observe that $d[\mathsf{stub}]$ has a stub-instance of type $\mathsf{stub} \quad \tau/\mathcal{S}$, with $\mathcal{S} \equiv B \vee (A \supset B) \Rightarrow A \supset B$, and d' has endsequent $\mathcal{S}' \equiv B \Rightarrow A \supset B$.

Since $\mathcal{S}' \quad A \supset B \Rightarrow A \supset B/\mathcal{S}$ is an instance of $L\vee$, $d[\mathsf{stub}]$ and d' are compatible. The derivation $d[\mathsf{stub}] \leftarrow\!\dashv d'$ is identical to that obtained in Example 3, although here the binding rule is $L\vee$.

DEFINITION 5. Let τ be a stub-derivation and G a formula multiset. The antecedent operator $\oplus$: stub-derivation $\times$ formula multiset $\mapsto$ stub-derivation is defined as follows:

Case $G = \langle\rangle$: let $\tau \oplus G = \tau$

Case $G \neq \langle\rangle$: define $\tau \oplus G$ recursively on τ as follows

1. initial sequent: $(A \Rightarrow A) \oplus G = (A \Rightarrow A/^{LW^*(G)} A, G \Rightarrow A)$
2. stub-instance:
   (a) $(\mathsf{stub}/X \Rightarrow Y) \oplus G = (\mathsf{stub}/X, G \Rightarrow Y)$
   (b) $(\mathsf{stub} \quad \pi/X \Rightarrow Y) \oplus G = (\mathsf{stub} \quad \pi \oplus G/X, G \Rightarrow Y)$
   (c) $(\pi \quad \mathsf{stub}/X \Rightarrow Y) \oplus G = (\pi \oplus G \quad \mathsf{stub}/X, G \Rightarrow Y)$
3. unary non-GLR rule: $(\pi/X \Rightarrow Y) \oplus G = (\pi \oplus G/X, G \Rightarrow Y)$
4. GLR rule: $(\pi/^{GLR} X \Rightarrow Y) \oplus G = (\pi/^{GLR} X \Rightarrow Y)/^{LW^*(G)} X, G \Rightarrow Y$
5. binary additive rule: $(\pi_1\ \pi_2/X \Rightarrow Y) \oplus G = (\pi_1 \oplus G\ \pi_2 \oplus G/X, G \Rightarrow Y)$
6. cut-rule: $(\pi_1 \quad \pi_2/^{cut} X \Rightarrow Y) \oplus G = (\pi_1 \oplus G \quad \pi_2/^{cut} X, G \Rightarrow Y)$.

That $\oplus$ maps into the set of stub-derivations follows by inspection of the definition. Notice that the recursion terminates at an initial sequent, stub-instance or a GLR rule. The operator $\oplus$ will bind stronger that $\leftarrow\!\dashv$.

LEMMA 6. *If d is a stub-derivation and G is a formula multiset, then $d \oplus G$ is a stub-derivation. Furthermore, if d is in fact an sstub-derivation $d[\mathsf{stub}]$, then $d[\mathsf{stub}] \oplus G$ is an sstub-derivation.*

Proof. The result follows immediately from Definition 5. ■

EXAMPLE 7. Refer to the sstub-derivation $d[\mathsf{stub}]$ in Example 3. If G is a non-empty formula multiset, then $d[\mathsf{stub}] \oplus G$ is the stub-derivation:

$$\dfrac{\dfrac{\mathsf{stub}}{B, G \Rightarrow A \supset B} \qquad \dfrac{A \supset B \Rightarrow A \supset B}{A \supset B, G \Rightarrow A \supset B}\ LW^*(G)}{B \vee (A \supset B), G \Rightarrow A \supset B}\ L\vee$$

Form of annotated derivation δ	$\Phi_{\Box B}[\delta]$
$(\Box B)^* \Rightarrow \Box B$	$(\Box B)^\circ \Rightarrow \Box B$
$\dfrac{\dfrac{\{\pi\}_1^r}{G, (\Box B)^{n-1} \Rightarrow H}}{G, (\Box B)^{*n} \Rightarrow H}\ LW(\Box B)$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G, (\Box B)^{*n-1} \Rightarrow H}\right]}{G, (\Box B)^\circ, (\Box B)^{*n-1} \Rightarrow H}\ LW(\Box B)$
$\dfrac{\dfrac{\{\pi\}_1^r}{G, (\Box B)^{n+1} \Rightarrow H}}{G, (\Box B)^{*n} \Rightarrow H}\ LC(\Box B)$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G, (\Box B)^{*n+1} \Rightarrow H}\right]}{G, (\Box B)^{*n} \Rightarrow H}\ LC(\Box B)$
$\dfrac{\dfrac{\{\pi\}_1^r}{G', (\Box B)^{n} \Rightarrow H'}}{G, (\Box B)^{*n} \Rightarrow H}\ \rho \neq GLR$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G', (\Box B)^{*n} \Rightarrow H'}\right]}{G, (\Box B)^{*n} \Rightarrow H}\ \rho$
$\dfrac{\dfrac{\{\pi\}_1^r}{\Box G, G, (\Box B)^{n}, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{*n} \Rightarrow \Box A}\ \mathrm{GLR}$	$\dfrac{\dfrac{\{\pi\}_1^r}{\Box G, G, (\Box B)^{n}, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\ \mathrm{GLR}$
$\dfrac{\dfrac{\{\pi\}_1^r}{G', (\Box B)^{n} \Rightarrow H'} \quad \dfrac{\{\pi'\}_1^s}{G'', (\Box B)^{n} \Rightarrow H''}}{G, (\Box B)^{*n} \Rightarrow H}\ \rho \neq cut$	$\dfrac{\Phi_{\Box B}\left[\dfrac{\{\pi\}_1^r}{G', (\Box B)^{*n} \Rightarrow H'}\right] \quad \Phi_{\Box B}\left[\dfrac{\{\pi'\}_1^s}{G'', (\Box B)^{*n} \Rightarrow H''}\right]}{G, (\Box B)^{*n} \Rightarrow H}\ \rho$
antecedent of $ES(\delta)$ does not contain a $(\Box B)^*$ formula occurrence	δ

Table 2. Definition of $\Phi_{\Box B}$. Multisets G and $\Box G$ contain no occurrences of annotated formulae. The function is defined on the class of cut-free annotated derivations.

By observation, we can confirm that $d[\mathsf{stub}] \oplus G$ is a sstub-derivation as predicted by Lemma 6. Notice that $d[\mathsf{stub}] \oplus G$ and d' (from Example 3) are not compatible, because there is no logical or structural inference rule that can take us from the premise sequent $A, B \Rightarrow B$ to the conclusion sequent $B, G \Rightarrow A \supset B$.

Definition 5 can be extended in the obvious way to apply to annotated stub-derivations. Then it is easily verified that Lemma 6 holds under the substitution of "annotated (s)stub-derivation" for "(s)stub-derivation" in the statement of the lemma.

Cut-elimination often involves tracing the "parametric ancestors" of the cut-formula. The following definition uses the symbols $\circ$ and $*$ as annotations to help trace the parametric ancestors.

DEFINITION 8. ($f_C[\cdot]$: annotated derivation wrt C).
Let τ be a cut-free derivation with endsequent $X \Rightarrow Y$, and C a formula.

1. if C is not boxed then let $f_C[\tau] = \tau$.
2. if C is boxed ($C \equiv \Box B$) and $\Box B \notin X$ then let $f_{\Box B}[\tau] = \tau$.
3. if C is boxed ($C \equiv \Box B$) and $\Box B \in X$. Then τ must be a derivation of the form $\Box B \Rightarrow \Box B$ or $\{\pi\}_1^r/X', \Box B \Rightarrow Y$.

   Set $f_{\Box B}[\tau]$ as $\Phi_{\Box B}[(\Box B)^* \Rightarrow \Box B)]$ or $\Phi_{\Box B}[\{\pi\}_1^r/X', (\Box B)^* \Rightarrow Y]$ respectively, where $\Phi_{\Box B}$ is defined on the class of cut-free annotated derivations as shown in Table 2.

Observe that the annotation operator $f_C[\cdot]$ is a total function mapping derivations to annotated derivations.

REMARK 9. Let τ be a derivation with endsequent $X \Rightarrow Y$. If $\Box B \in X$ then the formula occurrences $(\Box B)^\circ$ and $(\Box B)^*$ in $f_{\Box B}[\tau]$ are each called a *parametric ancestor* of the formula occurrence $\Box B \in X$ in the endsequent. Intuitively, the annotation $\circ$ denotes the final parametric ancestor when tracing ancestors upwards. That is, the $\Box B$ is introduced at that point.

DEFINITION 10. Define $\partial^\circ(B, \tau)$ for a formula B and an annotated derivation τ, as the number of instances of the GLR rule in τ whose conclusion contains an occurrence of the annotated formula B° in the antecedent.

LEMMA 11. *Let $d[\mathsf{stub}]$ be an annotated sstub-derivation and G a formula multiset. Then*

(a) $\partial^\circ(B, d[\mathsf{stub}] \oplus G) = \partial^\circ(B, d[\mathsf{stub}])$

(b) Let d' be a derivation such that $d[\mathsf{stub}]$ and d' are compatible. Then $\partial^\circ(B, d[\mathsf{stub}] \hookleftarrow d') = \partial^\circ(B, d[\mathsf{stub}]) + \partial^\circ(B, d')$.

Proof.

(a) Because $\partial^\circ(B, \cdot)$ counts the number of instances of the GLR rule with conclusion sequents containing the formula occurrence B°, the result is an immediate consequence of the fact that $\oplus$ does not introduce formulae into the conclusion sequent of an instance of the GLR rule (see Definition 5(4)).

(b) By the definition of compatibility, the binding rule for $d[\mathsf{stub}]$ and d' cannot be GLR. Thus if an instance ρ of the GLR rule appears in $d[\mathsf{stub}] \hookleftarrow d'$, then ρ must appear in one of $d[\mathsf{stub}]$ or d'. Also, if an instance ρ of the GLR rule appears in either $d[\mathsf{stub}]$ or d', then ρ must appear in $d[\mathsf{stub}] \hookleftarrow d'$. The result follows immediately.

■

REMARK 12. Lemma 11(a) holds even if G contains decorated formulae.

DEFINITION 13 (width). Let cut_0 be a topmost cut as shown below:

$$\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, B}\,\rho \qquad \dfrac{\{\sigma\}_1^s}{B, U \Rightarrow W}}{X, U \Rightarrow Y, W}\,cut_0$$

Then, the width of cut_0 is defined as:

$$width(cut_0) = \begin{cases} \partial^\circ(B, f_B[\pi_1]) & \text{if } \rho = GLR \text{ (so } \{\pi\}_1^r = \pi_1) \\ GLR(2, \{\pi\}_1^r / X \Rightarrow Y, B) & \text{otherwise} \end{cases}$$

REMARK 14.

(i) The width has been defined only for a topmost cut as this context is sufficient for our purposes.

(ii) $width(cut_0)$ is independent of the right premise derivation of cut_0.

EXAMPLE 15. Let us calculate $width(cut_0)$ in the following:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box C, C, \Box\Box B, \Box B, \Box B \Rightarrow B}}{\Box C, \Box\Box B \Rightarrow \Box B}\,GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{\Box D \Rightarrow \Box B}}{\Box D, \Box\Box B \Rightarrow \Box B}\,LW}{\Box C \vee \Box D, \Box\Box B \Rightarrow \Box B}\,L\vee}{\Box(\Box C \vee \Box D), \Box C \vee \Box D, \Box\Box B \Rightarrow \Box B}\,LW}{\Box(\Box C \vee \Box D) \Rightarrow \Box\Box B}\,GLR \qquad \Box\Box B, U \Rightarrow W}{\Box(\Box C \vee \Box D), U \Rightarrow W}\,cut_0$$

Writing the left premise derivation of cut_0 as $\mu / \Box(\Box C \vee \Box D) \Rightarrow \Box\Box B$, we get $width(cut_0) = \partial^\circ(\Box\Box B, f_{\Box\Box B}[\mu])$ where $f_{\Box\Box B}[\mu]$ is

$$\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box C, C, \Box\Box B, \Box B, \Box B \Rightarrow B}}{\Box C, (\Box\Box B)^\circ \Rightarrow \Box B}\,GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{\Box D \Rightarrow \Box B}}{\Box D, (\Box\Box B)^\circ \Rightarrow \Box B}\,LW}{\Box C \vee \Box D, (\Box\Box B)^* \Rightarrow \Box B}\,L\vee}{\Box(\Box C \vee \Box D), \Box C \vee \Box D, (\Box\Box B)^* \Rightarrow \Box B}\,LW$$

Because $f_{\Box\Box B}[\mu]$ contains only one GLR rule whose conclusion contains the formula occurrence $(\Box\Box B)^\circ$ in its antecedent, we have $width(cut_0) = 1$.

REMARK 16. Let μ be the left premise derivation of cut_0 from Definition 13. Valentini [16, pg 473] defines the width as the cardinality of $GLR^{(2)}$, where $GLR^{(2)}$ in our notation is the set of all instances ρ of GLR such that:

(a) ρ is a 2-ary GLR rule in μ, and

(b) B is the diagonal formula of every 1-ary GLR rule in μ below ρ, and

(c) B is not introduced by weakening below ρ.

Applying Valentini's original definition to the following derivation in GLS

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box\Box X, \Box X, \Box X, X, \Box\Box C, \Box C, \Box C \Rightarrow C}}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, GLR}{\Box\Box X, \Box X, \Box\Box C, \Box\Box C \Rightarrow \Box C}\, LW(\Box\Box C)}{\Box\Box X, \Box X, \Box\Box C \Rightarrow \Box C}\, LC(\Box\Box C)}{\Box\Box X \Rightarrow \Box\Box C}\, GLR \qquad \Box\Box C, U \Rightarrow W}{\Box\Box X, U \Rightarrow W}\, cut_0$$

we compute the width of cut_0 as 0 (due to condition *(c)*). Using the definition in this paper we have $width(cut_0) = 1$. Our definition considers the interplay of the weakening and contraction rules, and is required to obtain the cut-elimination result for GLS. In GLS_V however, there are no contraction rules so Valentini's original definition suffices.

Thus Moen is certainly justified in asking whether Valentini's arguments can be lifted to multiset-based sequents. However, we will see that Moen's claims about failure of cut-elimination in the new setting are incorrect.

4 Cut-elimination for GLS

We have the following decomposition lemma.

LEMMA 17. *Let τ be a cut-free derivation of the form $\{\pi\}_1^r/^{\rho} X, \Box B \Rightarrow Y$ and suppose that $\partial^\circ(\Box B, f_{\Box B}[\tau]) > 0$. If*

(i) $\rho = GLR$ then $f_{\Box B}[\tau] = \{\pi\}_1^r/^{GLR} X, (\Box B)^\circ \Rightarrow Y$.

(ii) $\rho \neq GLR$ then we can write the annotated derivation $f_{\Box B}[\tau]$ in the form $d[\mathsf{stub}] \hookleftarrow d'$ such that

$$\partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d').$$

Furthermore, denote the endsequent of d' as $U \Rightarrow W$. Then for any multiset M, and any derivation d'' with endsequent $U, M \Rightarrow W$, we have that $d[\mathsf{stub}] \oplus M$ and d'' are compatible.

Proof. If $\rho = GLR$ then it follows immediately from Definition 8 that $f_{\Box B}[\tau] = \{\pi\}_1^r/^{GLR} X, (\Box B)^\circ \Rightarrow Y$.

Now suppose that $\rho \neq GLR$. Because $\partial^\circ(\Box B, f_{\Box B}[\tau]) > 0$, we can write $f_{\Box B}[\tau]$ in the form below, where $n \geq 1$, and $\Box G$ contains no annotated formulae, and $\eta/X, (\Box B)^* \Rightarrow Y$ is an annotated sstub-derivation possibly containing branches:

$$\begin{array}{c} \dfrac{\dfrac{\{\pi'\}_1^s}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\, GLR_1 \\ \eta \\ X, (\Box B)^* \Rightarrow Y \end{array}$$

We can identify the annotated derivation $f_{\Box B}[\tau]$ with $d[\mathsf{stub}] \hookleftarrow d'$ where d' (below left) is an annotated derivation and $d[\mathsf{stub}]$ (below right) is an annotated sstub-derivation.

$$\dfrac{\dfrac{\{\pi'\}_1^s}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\,GLR \qquad\qquad \begin{matrix}\mathsf{stub}\\ \eta\\ X, (\Box B)^* \Rightarrow Y\end{matrix}$$

From Lemma 11(b) we have

$$\partial^\circ(\Box B, f_{\Box B}[\tau]) = \partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d').$$

Write the endsequent of d' as $U \Rightarrow W$. Observe that GLR_1 must be a 1-ary GLR rule in $f_{\Box B}[\tau]$. If this were not the case, the antecedent of the conclusion of GLR_1 could not contain occurrences of B°. Thus the path (through η) between the leaf stub in $d[\mathsf{stub}]$ and the endsequent $X, (\Box B)^* \Rightarrow Y$ of $d[\mathsf{stub}]$ contains no GLR rule instances. From Definition 5 and the compatibility of $d[\mathsf{stub}]$ and d', for any multiset M and any derivation d'' with endsequent $U, M \Rightarrow W$, it follows that $d[\mathsf{stub}] \oplus M$ and d'' are compatible. ■

DEFINITION 18 (rank of a cut). For a topmost cut cut_0 the *rank* $rk(cut_0)$ is the triple (d, n, h) where d is the degree of the cut-formula, n is the width of cut_0, and h is the cut-height of cut_0.

LEMMA 19. *Let τ be the following derivation where cut_0 is a topmost cut:*

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\{\sigma\}_1^s}{\Box B, U \Rightarrow W}}{\Box X, U \Rightarrow W}\,cut_0$$

and suppose ($\star$): for any derivation δ, every topmost cut in δ with rank $<$ $rk(cut_0)$ is eliminable.

Then there is a transformed cut-free derivation τ' of $X, \Box X \Rightarrow B$.

Proof. Let μ denote the subderivation $\{\pi\}_1^r/\Box X, X, \Box B \Rightarrow B$ of τ.

Case width(cut$_0$) = 0 : Hence $\partial^\circ(\Box B, f_{\Box B}[\mu]) = 0$. Then the annotated derivation $f_{\Box B}(\mu)$ must have final parametric ancestors of the form $(\Box B)^\circ \Rightarrow \Box B$ or $X' \Rightarrow Y'/^{LW(\Box B)}X', (\Box B)^\circ \Rightarrow Y'$ only.

Let $\Box B^{(*|\circ)}$ stand for an annotated occurrence of $\Box B$ where the annotation is not known. Consider the substitution $(f_{\Box B}[\mu])\{\Box B^{(*|\circ)} := \Box X\}$ obtained by replacing every occurrence

1. of $(\Box B)^*$ with $\Box X$, and
2. of $(\Box B)^\circ \Rightarrow \Box B$ with a derivation of $\Box X \Rightarrow \Box B$ (the left premise derivation of cut_0), and
3. of $\dfrac{X' \Rightarrow Y'}{X', (\Box B)^\circ \Rightarrow Y'}\,LW(\Box B)$ with $\dfrac{X' \Rightarrow Y'}{X', \Box X \Rightarrow Y'}\,LW^*(\Box X)$

As the endsequent of $f_{\Box B}[\mu]$ was $\Box X, X, (\Box B)^* \Rightarrow B$ we have that $(f_{\Box B}[\mu])\{\Box B^{(*|\circ)} := \Box X\}$ is a cut-free derivation of $\Box X, X, \Box X \Rightarrow B$. Applying repeated left contraction gives a cut-free derivation of $\Box X, X \Rightarrow B$.

Case width(cut$_0$) > 0 : Hence $\partial^\circ(\Box B, f_{\Box B}[\mu]) > 0$. First suppose that the last rule in μ is GLR. Then μ must be of the form:

$$\dfrac{\dfrac{\{\pi'\}_1^s}{\Box\Box X', \Box X', \Box X', X', \Box\Box A, \Box A, \Box A \Rightarrow A}}{\Box\Box X', \Box X', \Box\Box A \Rightarrow \Box A}\,GLR$$

where $X = \Box X'$ and $B \equiv \Box A$.

Then the following is a derivation of $\Box X, X \Rightarrow B$, with $deg(cut_1) = deg(cut_0)$ and $width(cut_1) = 0 < width(cut_0)$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Box A \Rightarrow \Box A}{\Box A, A, \Box\Box A \Rightarrow \Box A}\,LW^*(A, \Box\Box A)}{\Box A \Rightarrow \Box\Box A}\,GLR \qquad \dfrac{\{\pi'\}_1^s}{\Box\Box X', \Box X', \Box X', X', \Box\Box A, \Box A, \Box A \Rightarrow A}}{\Box A, \Box\Box X', \Box X', \Box X', X', \Box A, \Box A \Rightarrow A}\,cut_1}{\Box\Box X', \Box X', \Box X', X', \Box A \Rightarrow A}\,LC^*(\Box A)}{\Box\Box X', \Box X' \Rightarrow \Box A}\,GLR$$

The required derivation is obtained by using $(\star)$ to eliminate cut_1.

If the last rule in μ is not GLR, by Lemma 17 we can write $f_{\Box B}[\mu]$ as $d[\mathsf{stub}] \hookleftarrow d'$, where d' and $d[\mathsf{stub}]$ are respectively:

$$\dfrac{\dfrac{\{\pi'\}_1^t}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\,GLR \qquad\qquad \begin{array}{c}\mathsf{stub}\\ \eta\\ \Box X, X, (\Box B)^* \Rightarrow B\end{array}$$

where $n \geq 1$, and $\Box G$ does not contain annotated formulae, and

$$\partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') = \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d').$$

Let d'' be the annotated derivation

$$\dfrac{\Box A \Rightarrow \Box A}{A, \Box A, \Box G, (\Box B)^{\circ n} \Rightarrow \Box A}\,LW^*(A, \Box G, (\Box B)^n)$$

Then $d[\mathsf{stub}] \oplus \langle A, \Box A\rangle$ and d'' are compatible (Lemma 17). Note that $\partial^\circ(\Box B, d') = 1$ and $\partial^\circ(\Box B, d'') = 0$. Let Λ_{11} be the derivation:

$$\dfrac{\dfrac{\dfrac{\lfloor d[\mathsf{stub}] \oplus \langle A, \Box A\rangle \hookleftarrow d''\rfloor}{\Box A, \Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box A, \Box X, \Box X, X \Rightarrow B}\,cut_1}{\Box A, \Box X, X \Rightarrow B}\,LC^*$$

and Λ_{12} the derivation

$$\dfrac{\dfrac{\dfrac{\lfloor d[\mathsf{stub}] \oplus \langle A, \Box A\rangle \hookleftarrow d''\rfloor}{\Box A, \Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\{\pi'\}_1^t}{\Box G, G, (\Box B)^n, B^n, \Box A \Rightarrow A}}{\Box A, \Box A, \Box X, \Box G, G, (\Box B)^{n-1}, B^n \Rightarrow A}\,cut_2}{\Box A, \Box X, \Box G, G, (\Box B)^{n-1}, B^n \Rightarrow A}\,LC$$

Consider the derivation Λ_1:

$$\dfrac{\dfrac{\dfrac{\Lambda_{11} \qquad \Lambda_{12}}{\Box A, \Box X, X, \Box A, \Box X, \Box G, G, (\Box B)^{n-1}, B^{n-1} \Rightarrow A}\,cut_3}{\Box A, \Box X, X, \Box G, G, (\Box B)^{n-1}, B^{n-1} \Rightarrow A}\,LC^*}{\Box X, \Box G, (\Box B)^{n-1} \Rightarrow \Box A}\,GLR$$

For $i \in \{1,2\}$, observe that $deg(cut_i) = deg(cut_0)$. Furthermore,

$$\begin{aligned}
width(cut_i) &= \partial^\circ(\Box B, f_{\Box B}(\lfloor d[\mathsf{stub}] \oplus \langle A, \Box A\rangle \hookleftarrow d''\rfloor)) && \text{Def. of width}\\
&= \partial^\circ(\Box B, d[\mathsf{stub}] \oplus \langle A, \Box A\rangle \hookleftarrow d'') && \text{By inspection}\\
&= \partial^\circ(\Box B, d \oplus \langle A, \Box A\rangle[\mathsf{stub}]) + \partial^\circ(\Box B, d'') && \text{Lemma 11(b)}\\
&< \partial^\circ(\Box B, d \oplus \langle A, \Box A\rangle[\mathsf{stub}]) + \partial^\circ(\Box B, d') &&\\
&= \partial^\circ(\Box B, d[\mathsf{stub}]) + \partial^\circ(\Box B, d') && \text{Lemma 11(a)}\\
&= \partial^\circ(\Box B, d[\mathsf{stub}] \hookleftarrow d') && \text{Lemma 11(b)}\\
&= width(cut_0)
\end{aligned}$$

Because $deg(cut_i) = deg(cut_0)$ and the premises of both cut_1 and cut_2 are cut-free, by appealing twice to $(\star)$ we can in turn eliminate cut_1 and cut_2. In the resulting derivation, since $deg(cut_3) < deg(cut_0)$ we can eliminate cut_3 by $(\star)$. We thus obtain a cut-free derivation Λ_2 of $\Box X, \Box G, (\Box B)^{n-1} \Rightarrow \Box A$.

Let Λ_3 be the annotated derivation

$$\dfrac{\Lambda_2}{\Box X, \Box G, (\Box B)^n \Rightarrow \Box A}\,LW(\Box B)$$

Clearly $\partial^\circ(\Box B, \Lambda_3) = 0$. Furthermore, by Lemma 17, $d[\mathsf{stub}] \oplus \Box X$ and Λ_3 are compatible. Recall that $\lfloor\cdot\rfloor$ is the forgetful map. The endsequent of $\lfloor(d[\mathsf{stub}] \oplus X) \hookleftarrow \Lambda_3\rfloor$ is thus $\Box X, \Box X, X, \Box B \Rightarrow B$. Now consider the derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{\lfloor(d[\mathsf{stub}] \oplus \Box X) \hookleftarrow \Lambda_3\rfloor}{\Box B, \Box X, X \Rightarrow B}\,LC^*(\Box X)}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{X, \Box X, \Box X \Rightarrow B}\,cut_4(\Box B)}{X, \Box X \Rightarrow B}\,LC^*(\Box X)$$

By a similar calculation to the above we obtain $width(cut_4) < width(cut_0)$. Because $deg(cut_4) = deg(cut_0)$ and the premises of cut_4 are cut-free, appealing to $(\star)$ we can eliminate cut_4. We thus obtain a cut-free derivation of $X, \Box X \Rightarrow B$ as required. ■

Without loss of generality it suffices to consider a derivation concluded by a cut-rule with cut-free premise derivations.

THEOREM 20 (Cut-elimination). *Let τ be a derivation concluded by an instance cut_0 of the cut-rule with cut-free premise derivations. Then there is a transformed cut-free derivation τ' with identical endsequent.*

Proof. Induction on the rank (d, n, h) of cut_0 under the standard lexicographic ordering. We say that the cut-formula is *left principal* if an occurrence of the cut-formula in the succedent of the left premise is a principal formula. The term *right principal* is defined analogously. This follows standard practice.

1 Cut with an initial sequent as premise. This is the base case. The transformations are standard (see [9],[15]).

2 Cut with neither premise an initial sequent.

(a) Cut-formula is left and right principal.

First suppose that the cut-formula is boxed. There are five possibilities:

(i) the cut-formula is left and right principal by the GLR rule. The derivation must then be in SNF:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{\Box B, \Box U, B, U, \Box C \Rightarrow C}}{\Box B, \Box U \Rightarrow \Box C}\,GLR}{\Box X, \Box U \Rightarrow \Box C}\,cut_0$$

The induction hypothesis implies that for any derivation δ, any topmost cut in δ with rank $< rank(cut_0)$ is eliminable. This is precisely condition $(\star)$ in Lemma 19. Hence we can obtain a cut-free derivation of $\Box X, X \Rightarrow B$. Consider the derivation

$$\dfrac{\dfrac{\Box X, X \Rightarrow B \qquad \dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\{\sigma\}_1^s}{\Box B, \Box U, B, U, \Box C \Rightarrow C}}{\Box X, \Box U, B, U, \Box C \Rightarrow C}\,cut_1}{\Box X, X, \Box X, \Box U, U, \Box C \Rightarrow C}\,cut_2}{\Box X, \Box U \Rightarrow \Box C}\,GLR$$

Observe that $rk(cut_1) = (d, n, h-1)$. By the induction hypothesis we can eliminate cut_1. In the resulting derivation, since $deg(cut_2) < d$, the result follows from another application of the induction hypothesis.

(ii) the cut-formula $\Box B$ is left principal by the GLR rule and right principal by $LC(\Box B)$.

Then τ is as below where both premises of cut_0 are cut-free and $m \geq 0$:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{(\Box B)^{m+2}, U \Rightarrow W}\,\rho}{\Box B, U \Rightarrow W}\,LC^{m+1}(\Box B)}{\Box X, U \Rightarrow W}\,cut_0$$

In general ρ need not be the GLR rule. However if $\rho \neq GLR$ then either (1) $\rho = LW(\Box B)$ and we delete ρ and the $LC(\Box B)$ rule that follows, or (2) $\Box B$ is not principal by ρ.

In the former case the result is immediate. In the latter case the result is obtained by applying ρ *after* cut_0 and invoking the induction hypothesis. Observe that this is possible even if ρ is a binary rule.

If $\rho = GLR$ it follows that $U \equiv \Box V$ and $W \equiv \Box C$ for some multiset V and formula C, and $s = 1$ and $\sigma_1 \equiv \{\sigma'\}_1^{s'}/(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C$. Thus τ must be of the form

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{\Box X, X, \Box B \Rightarrow B}}{\Box X \Rightarrow \Box B}\,GLR \qquad \dfrac{\dfrac{\dfrac{\{\sigma'\}_1^{s'}}{(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C}}{(\Box B)^{m+2}, \Box V \Rightarrow \Box C}\,\rho = GLR}{\Box B, \Box V \Rightarrow \Box C}\,LC^{m+1}(\Box B)}{\Box X, \Box V \Rightarrow \Box C}\,cut_0$$

A derivation of $\Box X, X \Rightarrow B$ is obtained as in (i) using Lemma 19. Consider the derivation:

$$\dfrac{\dfrac{\dfrac{\Box X, X \Rightarrow B \qquad \dfrac{\dfrac{\Box X \Rightarrow \Box B \qquad \dfrac{\dfrac{\{\sigma'\}_1^{s'}}{(\Box B)^{m+2}, B^{m+2}, \Box V, V, \Box C \Rightarrow C}}{\Box B, B^{m+2}, \Box V, V, \Box C \Rightarrow C}\,LC^{m+1}(\Box B)}{\Box X, B^{m+2}, \Box V, V, \Box C \Rightarrow C}\,cut_1}{\Box X, B, \Box V, V, \Box C \Rightarrow C}\,LC^{m+1}(B)}{\Box X, X, \Box X, \Box V, V, \Box C \Rightarrow C}\,cut_2}{\Box X, X, \Box V, V, \Box C \Rightarrow C}\,LC^*}{\Box X, \Box V \Rightarrow \Box C}\,GLR$$

Now cut_1 has identical degree and width compared to cut_0, and smaller cut-height. Hence, we can eliminate cut_1 using the induction hypothesis. In the resulting derivation $deg(cut_2) < d$ so the result follows from the induction hypothesis.

(iii) the cut-formula $\Box B$ is left principal by $RC(\Box B)$ and right principal by the GLR rule.

Then τ has the following form where both premises of cut_0 are cut-free:

$$\dfrac{\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, \Box B, \Box B}}{X \Rightarrow Y, \Box B}\,RC_1 \qquad \dfrac{\sigma}{\Box B, \Box U \Rightarrow \Box C}\,GLR}{X, \Box U \Rightarrow Y, \Box C}\,cut_0$$

Because the conclusion of (any) GLR rule has a exactly one formula in the succedent, it follows that at least one of the $\Box B$ formula occurrences in the succedent of the premise of RC_1 can be traced upwards in $\{\pi\}_1^r$ to $RW(\Box B)$ rule application(s). In particular, when tracing upwards, it is impossible to encounter a GLR rule application *before* encountering a $RW(\Box B)$ rule application. Deleting these $RW(\Box B)$ rule applications and the RC_1 contraction rule certainly preserves the derivation structure because all the binary rules excluding the cut-rule are additive. This new derivation contains a single instance of cut with identical degree of cut-formula and reduced cut-height compared to cut_0. Furthermore, observe that it must be the case that the width is $\leq n$. The result follows from the induction hypothesis.

If the calculus uses multiplicative binary rules instead, the result still holds, although the transformations are slightly more complicated.

In each instance, the proof can be formalised using an annotation function similar in structure to $f_{\Box B}$. We omit the details.

(iv) the cut-formula $\Box B$ is left and right principal by $RC(\Box B)$ and $LC(\Box B)$ respectively.

A combination of the strategies in (ii) and (iii) suffice.

(v) the cut-formula $\Box B$ is either left or right principal by $RW(\Box B)$ or $LW(\Box B)$ respectively.

Trivial.

When the cut-formula is not boxed and the cut-formula is left and right principal by the respective left and right introduction rule the transformations are standard (see [9],[15] for example) — derivation τ is transformed to a derivation τ' containing cuts $\{cut_i\}_{i\geq 1}$ on strictly smaller cut-formulae (i.e. $deg(cut_i) < d$ for $i \geq 1$).

If the cut-formula is right principal by $LC(B)$ then τ has the form below where B is principal by ρ:

$$\dfrac{\dfrac{\{\pi\}_1^r}{X \Rightarrow Y, B}\,\rho \qquad \dfrac{\dfrac{\{\sigma\}_1^s}{B, B, U \Rightarrow W}}{B, U \Rightarrow W}\,LC(B)}{X, U \Rightarrow Y, W}\,cut_0$$

We must have $\rho \neq GLR$. This is the well-known case of contractions above cut that occurs in cut-elimination for LK.

There are several proofs of cut-elimination avoiding Gentzen's multicut rule, for classical sequent calculi, appearing in the literature (for example, see [10],[3],[1]). We adapt the transformations proposed by von Plato [10] for the classical calculus $G0c$. Although all the binary rules in $G0c$ are multiplicative, and all the binary rules in GLS are additive, the same argument can be lifted here.

That argument in [10] relies on invertibility of all logical rules in $G0c$. Invertibility is not required to be height-preserving. A similar result holds for GLS too. We omit the details.

The cases corresponding to (iii)-(v) can be dealt with similarly.

(b) Cut-formula is left principal only.

(c) Cut-formula neither left nor right principal.

We analyse the last inference rule in the *right (left)* premise derivation of cut_0. The standard transformations suffice here (see [9],[15] for example). In particular, observe that for any instance cut_1 of the cut-rule appearing in a transformed derivation, it must be the case that $width(cut_1) \leq n$. ■

REMARK 21. In general, it is possible for the width of lower cuts to increase under the cut-elimination transformations. For example, consider some transformation which reduces some topmost cut instance cut_b (for "before") to the derivation below containing the cut instance cut_a (for "after") where $\{\pi\}_1^r$ and $\{\sigma\}_1^s$ need not be cut-free:

$$\dfrac{\{\pi\}_1^r \qquad \{\sigma\}_1^s}{G \Rightarrow H}\,cut_a$$

The cut-elimination transformations which ultimately turn cut_a into a topmost cut may produce a derivation where $width(cut_a) > width(cut_b)$. However, we have seen in the proof of Lemma 19 that $width(cut_4)$ does not increase despite any reductions above it. This is because the cut_4 in that proof is 'shielded' by the GLR instance concluding Λ_1. This shielding is crucial for the success of the proof.

5 Moen's Val-II(core) is not Valentini's reduction

We have carefully examined Moen's slides titled "The proposed algorithms for eliminating cuts in the provability calculus GLS do not terminate" [7].

Moen sets out to reduce a cut in SNF using the transformation he titles Val-II(core). Moen claims that Val-II(core) is the "…core of Valentini's reduction" [7]. Yet Val-II(core) does not appear in [16]. However it appears in [11, page 322] with the comment "this reduction is not sufficient".

Thus Moen is incorrect in claiming that he has demonstrated that Valentini's algorithm does not terminate — Moen is using the wrong algorithm. In fact, for his concrete derivation ϵ, the width of the cut-formula is 1 so the reduction is immediate. Applying the base case transformations, and then the classical transformations, we obtained a cut-free derivation of the end-sequent of ϵ.

6 Conclusion

We have resolved the issue surrounding the use of Valentini's arguments for cut-elimination in a multiset setting for GL. In order to formally define the measure width, we formalised the notion of 'tracing up' a derivation (i.e. identifying the parametric ancestors) via a constructive function. This constructive function can be used to aid the formalisation of various other notions in proof theory.

BIBLIOGRAPHY

[1] K. Bimbó. $LE^t_{\rightarrow}$, $LR^{\circ}_{\underset{\sim}{\wedge}}$, LK and cut-free proofs. *Journal of Philosophical Logic*, 36:557–570, 2007.

[2] M. Borga. On Some Proof Theoretical Properties of the Modal Logic GL. *Studia Logica*, 42:453–459, 1983.

[3] M. Borisavljević, K. Došen and Z. Petrić. On permuting cut with contraction. *Math. Struct. in Comp. Science*, 10:99–136, 2000.

[4] G. Gentzen. The Collected Papers of Gerhard Gentzen, ed. M. Szabo.

[5] D. Leivant. On the Proof Theory of the Modal Logic for Arithmetic Provability. *Journal of Symbolic Logic*, 46:531–538, 1981.

[6] G. Mints. Cut elimination for provability logic. Collegium Logicum 2005.

[7] A. Moen. The proposed algorithms for eliminating cuts in the provability calculus GLS do not terminate. *NWPT 2001*, Norwegian Computing Center, 2001-12-10. http://publ.nr.no/3411

[8] S. Negri. Proof Analysis in Modal Logic. *Journal of Philosophical Logic*, 34:507–544, 2005.

[9] S. Negri and J. von Plato. *Structural Proof Theory.* CUP, 2001.

[10] J. von Plato. A proof of Gentzen's *Hauptsatz* without multicut. *Archive of Mathematical Logic*, 40:9–18, 2001.

[11] G. Sambin and S. Valentini. The Modal Logic of Provability. The Sequential Approach. *Journal of Philosophical Logic*, 11:311–342, 1982.

[12] K. Sasaki. Löb's Axiom and Cut-elimination Theorem. *Journal of the Nanzan Academic Society Math. Sci. and Information Engineering*, 1:91–98, 2001.

[13] V. Švejdar. On Provability Logic. *Nordic Journal of Philosophical Logic*, 4:95–116, 2000.

[14] R.M. Solovay. Provability Interpretations of Modal Logic. *Israel Journal of Mathematics*, 25:287–304, 1976.

[15] A.S. Troesltra and H. Schwichtenberg. *Basic Proof Theory.* CUP, 2000.

[16] S. Valentini. The Modal Logic of Provability: Cut-elimination. *Journal of Philosophical Logic*, 12:471–476, 1983.

Rajeev Goré
Computer Sciences Laboratory,
The Australian National University
Australia
`Rajeev.Gore@anu.edu.au`

Revantha Ramanayake
Computer Sciences Laboratory,
The Australian National University
Australia
`revantha@rsise.anu.edu.au`