# Cut-elimination theorem for second-order logic

(Full version)

April 24, 2013

## 1 Introduction

The result that the theorem of cut-elimination for pure predicative logic is formalizable in the theory $\Delta I_0 +$ superexp seems to be folklore. The proof of Takeuti in [**?**] is not readily formalizable in this theory because of an extra nested induction. The problem hinges on the fact that a (predicative) instantiation of a second-order quantification can arise through first-order formulas of arbitrary complexity.

In this note, we outline a proof strategy for cut-elimination in pure predicative logic with super-exponential increase in the length of the derivation. This proof has two steps: first, we eliminate all cuts whose cut-formula is second-order; then we invoke the usual cut-elimination result for first-order logic. Since the main ideas are the same as for the usual theorem of cut-elimination for first-order logic, we shall only emphasize the differences.

## 2 Cut-elimination for second-order logic

The syntax of second-order formulas is as usual. A formula with no second-order quantifiers (but possibly with free second-order variables) is a *first-order abstract*. A *sequent* is a pair $\langle \Gamma, \Delta \rangle$ of sets of formulas, written $\Gamma \vdash \Delta$, with intended semantics "if every formula in $\Gamma$ is true, then at least one formula in $\Delta$ is true".

The derivation rules for the second-order logic sequent calculus are summarized in Table 1. The following restrictions hold.

- In rules ($\forall R$) and ($\exists L$), $x$ is not among the free variables in $\Gamma \cup \Delta$.
- In rules ($\forall^2 R$) and ($\exists^2 L$), $R$ is not among the free second-order variables in $\Gamma \cup \Delta$.
- In rules ($\forall^2 L$) and ($\exists^2 R$), $\psi$ is a predicative first-order abstract with only one free variable.

In all rules but (Cut), the explicit formula in the conclusion is the rule's *principal formula*, whereas the explicit formula(s) in the premise(s) is (are) the *minor formula(s)*. In (Cut), $\varphi$ is the *cut formula*. The remaining

$$\frac{\Gamma \cap \Delta \neq \emptyset}{\Gamma \vdash \Delta} \text{ (Ax)}$$

$$\frac{\Gamma \vdash \Delta, \varphi}{\Gamma, \neg\varphi \vdash \Delta} \text{ ($\neg$ L)} \qquad \frac{\Gamma, \varphi \vdash \Delta}{\Gamma \vdash \Delta, \neg\varphi} \text{ ($\neg$ R)}$$

$$\frac{\Gamma, \varphi \vdash \Delta}{\Gamma, \varphi \wedge \psi \vdash \Delta} \text{ ($\wedge$ L$_1$)} \qquad \frac{\Gamma, \psi \vdash \Delta}{\Gamma, \varphi \wedge \psi \vdash \Delta} \text{ ($\wedge$ L$_2$)}$$

$$\frac{\Gamma \vdash \Delta, \varphi \qquad \Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \wedge \psi} \text{ ($\wedge$ R)}$$

$$\frac{\Gamma, \varphi \vdash \Delta \qquad \Gamma, \psi \vdash \Delta}{\Gamma, \varphi \vee \psi \vdash \Delta} \text{ ($\vee$ L)}$$

$$\frac{\Gamma \vdash \Delta, \varphi}{\Gamma \vdash \Delta, \varphi \vee \psi} \text{ ($\vee$ R$_1$)} \qquad \frac{\Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \vee \psi} \text{ ($\vee$ R$_2$)}$$

$$\frac{\Gamma \vdash \Delta, \varphi \qquad \Gamma, \psi \vdash \Delta}{\Gamma, \varphi \to \psi \vdash \Delta} \text{ ($\to$ L)}$$

$$\frac{\Gamma, \varphi \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \to \psi} \text{ ($\to$ R)}$$

$$\frac{\Gamma, \varphi[t/x] \vdash \Delta}{\Gamma, \forall x\varphi \vdash \Delta} \text{ ($\forall$ L)} \qquad \frac{\Gamma \vdash \Delta, \varphi}{\Gamma \vdash \Delta, \forall x\varphi} \text{ ($\forall$ R)}$$

$$\frac{\Gamma, \varphi \vdash \Delta}{\Gamma, \exists x\varphi \vdash \Delta} \text{ ($\exists$ L)} \qquad \frac{\Gamma \vdash \Delta, \varphi[t/x]}{\Gamma \vdash \Delta, \exists x\varphi} \text{ ($\exists$ R)}$$

$$\frac{\Gamma, \varphi[\psi/R] \vdash \Delta}{\Gamma, \forall R\varphi \vdash \Delta} \text{ ($\forall^2$ L)} \qquad \frac{\Gamma \vdash \Delta, \varphi}{\Gamma \vdash \Delta, \forall R\varphi} \text{ ($\forall^2$ R)}$$

$$\frac{\Gamma, \varphi \vdash \Delta}{\Gamma, \exists R\varphi \vdash \Delta} \text{ ($\exists^2$ L)} \qquad \frac{\Gamma \vdash \Delta, \varphi[\psi/R]}{\Gamma \vdash \Delta, \exists R\varphi} \text{ ($\exists^2$ R)}$$

$$\frac{\Gamma_1 \vdash \Delta_1, \varphi \qquad \Gamma_2, \varphi \vdash \Delta_2}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2} \text{ (Cut)}$$

Table 1: Derivation rules for the second-order sequent calculus. See the text for the side conditions on the rules involving quantifiers.

formulas in the rules are *side formulas*. The variable $x$ in rules ($\forall$ R) and ($\exists$ L) and the variable $R$ in rules ($\forall^2$ R) and ($\exists^2$ L) are said to be the *eigenvariables* of those rules.

**Definition 1.** *A* derivation *$d$ is a labeled tree where each node is labeled by either:*

- *an instance of (Ax) with no descendants;*
- *an instance of the conclusion of a unary rule with a single descendent labeled by the premise of that same rule;*
- *an instance of the conclusion of a binary rule with two descendants labeled by the premises of that same rule.*

Following tradition, we will write derivations upside-down (so that the root is at the bottom and the descendants of a node are above that node). We say that $d$ is a derivation of $\Gamma \vdash \Delta$, written $d \Vdash (\Gamma \vdash \Delta)$, if $d$ is a derivation and $d$'s root is labeled with $\Gamma \vdash \Delta$.

A sequent is said to be *normal* if all of its bound variables are distinct from all of its free variables. A derivation is said to be *normal* if (1) all of its sequents are normal and furthermore (2a) all second-order eigenvariables are distinct and (2b) every eigenvariable is used only above the node where it is an eigenvariable. Without loss of generality, we will assume that all derivations are normal, since it is trivial to change a derivation into a normal derivation.

The usual cut-elimination proof for first-order logic proceeds by induction on the cut-rank of derivations, replacing cuts of the highest rank with cuts on structurally smaller formulas. When we introduce second-order quantification, however, this no longer works in such a direct way, since eliminating a second-order quantifier typically produces structurally more complex formulas. For this reason, we split the cut-elimination proof in two steps. Both steps are similar, but they use different measures of complexity.

**Definition 2.** *The* size *of a first-order abstract $\varphi$, denoted $|\varphi|_1$, is defined inductively as usual.*

- $|A|_1 = 0$ *if $A$ is atomic.*
- $|\neg\varphi|_1 = |\forall x\varphi|_1 = |\exists x\varphi|_1 = |\varphi|_1 + 1$
- $|\varphi \vee \psi|_1 = |\varphi \wedge \psi|_1 = |\varphi \rightarrow \psi|_1 = \max(|\varphi|_1, |\psi|_1) + 1$

*Note that we do not include clauses for the second-order quantifiers, since these do not occur in first-order abstracts. We will often write $|\varphi|_1$ simply as $|\varphi|$.*

*The* second-order size *of a formula $\varphi$, denoted $|\varphi|_2$, is defined inductively as follows.*

- $|\psi|_2 = 0$ *if $\psi$ is a first-order abstract*
- $|\neg\varphi|_2 = |\forall x\varphi|_2 = |\exists x\varphi|_2 = |\varphi|_2 + 1$ *if $\varphi$ is* not *a first-order abstract (i.e. if $|\varphi|_2 > 0$)*
- $|\forall^2 R\varphi|_2 = |\exists^2 R\varphi|_2 = |\varphi|_2 + 1$
- $|\varphi \vee \psi|_2 = |\varphi \wedge \psi|_2 = |\varphi \rightarrow \psi|_2 = \max(|\varphi|_2, |\psi|_2) + 1$

The following result, which will have a key role in the proof of the Reduction Lemma below, is obtained by induction.

**Lemma 1.** *If $\psi$ is a first-order abstract, then $|\varphi[\psi/R]|_2 = |\varphi|_2$ for every formula $\varphi$.*

The length of a derivation is defined as usual.

**Definition 3.** *The* length *of a derivation $d$, $|d|$, is defined inductively.*

- *If $d$ consists of just an axiom, then $|d| = 0$.*
- *If $d$ ends with an application of a unary rule with subderivation $d'$, then $|d| = |d'| + 1$.*
- *If $d$ ends with an application of a binary rule with subderivations $d_1$ and $d_2$, then $|d| = \max(|d_1|, |d_2|)$.*

**Definition 4.** *The* second-order cut-rank *of a derivation $d$, $\rho_2(d)$ is defined inductively.*

- *If $d$ consists of just an axiom, then $\rho_2(d) = 0$.*
- *If $d$ ends with an application of a unary rule with subderivation $d'$, then $\rho_2(d) = \rho_2(d')$.*
- *If $d$ ends with an application of a binary rule other than (Cut) with subderivations $d_1$ and $d_2$, then $\rho_2(d) = \max(\rho_2(d_1), \rho_2(d_2))$.*
- *If $d$ ends with an application of (Cut) with subderivations $d_1$ and $d_2$ and cut-formula $\varphi$, then $\rho_2(d) = \max(\rho_2(d_1), \rho_2(d_2), |\varphi|_2)$.*

*If $\rho_2(d) = 0$, then the* first-order cut-rank *of $d$, $\rho_1(d)$, is also defined inductively.*

- *If $d$ consists of just an axiom, then $\rho_1(d) = 0$.*
- *If $d$ ends with an application of a unary rule with subderivation $d'$, then $\rho_1(d) = \rho_1(d')$.*
- *If $d$ ends with an application of a binary rule other than (Cut) with subderivations $d_1$ and $d_2$, then $\rho_1(d) = \max(\rho_1(d_1), \rho_1(d_2))$.*
- *If $d$ ends with an application of (Cut) with subderivations $d_1$ and $d_2$ and cut-formula $\varphi$, then $\rho_1(d) = \max(\rho_1(d_1), \rho_1(d_2), |\varphi|_1 + 1)$.*

*Note the important difference in the last clause of both definitions: a derivation with first-order cut-rank* $0$ *has no cuts, whereas a derivation with second-order cut-rank* $0$ *may have cuts with first-order abstracts as cut-formulas. Also note that the first-order cut-rank is only defined fot derivations with no second-order quantifiers in cut-formulas.*

The following results are directly proved by induction on $d$. In all of them, it is understood that the condition on $\rho_1(d)$ only applies if $\rho_2(d) = 0$.

**Lemma 2** (Weakening Lemma). *Suppose that* $d \Vdash (\Gamma \vdash \Delta)$, $\Gamma \subseteq \Gamma'$ *and* $\Delta \subseteq \Delta'$, *and that* $\Gamma' \cup \Delta'$ *is a normal sequent. Let* $d_{\Gamma,\Delta}^{\Gamma',\Delta'}$ *be obtained from* $d$ *by adding* $\Gamma' \setminus \Gamma$ *to the lefthandside of each sequent and* $\Delta' \setminus \Delta$ *to the righthandside of each sequent. Then:*

- $d_{\Gamma,\Delta}^{\Gamma',\Delta'} \Vdash (\Gamma' \vdash \Delta')$;
- $\left| d_{\Gamma,\Delta}^{\Gamma',\Delta'} \right| = |d|$;
- $\rho_i\left(d_{\Gamma,\Delta}^{\Gamma',\Delta'}\right) = \rho_i(d)$ *for* $i = 1, 2$.

**Lemma 3** (First-order Substitution Lemma). *Suppose that* $d \Vdash (\Gamma \vdash \Delta)$ *and that* $x$ *is not the eigenvariable of any application of* ($\forall$ *R*) *or* ($\exists$ *L*) *in* $d$. *Let* $d[s/x]$ *be obtained by replacing every occurrence of* $x$ *by a term* $s$ *in* $d$. *Then:*

- $d[s/x] \Vdash (\Gamma[s/x] \vdash \Delta[s/x])$;
- $|d[s/x]| = |d|$;
- $\rho_i(d[s/x]) = \rho_i(d)$ *for* $i = 1, 2$.

**Lemma 4** (Second-order Substitution Lemma). *Suppose that* $d \Vdash (\Gamma \vdash \Delta)$ *and that* $R$ *is not the eigenvariable of any application of* ($\forall^2$ *R*) *or* ($\exists^2$ *L*) *in* $d$. *Let* $d[\psi/R]$ *be obtained by replacing every occurrence of* $R$ *by the first-order abstract* $\psi$. *Then:*

- $d[\psi/R] \Vdash (\Gamma[\psi/R] \vdash \Delta[\psi/R])$;
- $|d[\psi/R]| = |d|$;
- $\rho_1(d[\psi/R]) \leq \rho_1(d) + |\varphi|_1$ *and* $\rho_2(d[\psi/R]) = \rho_2(d)$.

*Proof.* It is straightforward to check that $d[\psi/R] \Vdash (\Gamma[\psi/R] \vdash \Delta[\psi/R])$. The bound on $\rho_2(d[\psi/R])$ follows from Lemma 1, whereas the bound on $\rho_1(d[\psi/R])$ is proved by induction. $\square$

**Lemma 5** (Inversion Lemma).

1. *If* $d \Vdash (\Gamma, \neg\varphi \vdash \Delta)$, *then there is* $d_\varphi \Vdash (\Gamma \vdash \Delta, \varphi)$.
2. *If* $d \Vdash (\Gamma \vdash \Delta, \varphi \wedge \psi)$, *then there are* $d_\varphi \Vdash (\Gamma \vdash \Delta, \varphi)$ *and* $d_\psi \Vdash (\Gamma \vdash \Delta, \psi)$.
3. *If* $d \Vdash (\Gamma, \varphi \vee \psi \vdash \Delta)$, *then there are* $d_\varphi \Vdash (\Gamma, \varphi \vdash \Delta)$ *and* $d_\psi \Vdash (\Gamma, \psi \vdash \Delta)$.
4. *If* $d \Vdash (\Gamma, \varphi \to \psi \vdash \Delta)$, *then there are* $d_\varphi \Vdash (\Gamma \vdash \Delta, \varphi)$ *and* $d_\psi \Vdash (\Gamma, \psi \vdash \Delta)$.
5. *If* $d \Vdash (\Gamma \vdash \Delta, \forall x\varphi)$, *then there is* $d_\varphi \Vdash (\Gamma \vdash \Delta, \varphi)$.
6. *If* $d \Vdash (\Gamma, \exists x\varphi \vdash \Delta)$, *then there is* $d_\varphi \Vdash (\Gamma, \varphi \vdash \Delta)$.
7. *If* $d \Vdash (\Gamma \vdash \Delta, \forall R\varphi)$, *then there is* $d_\varphi \Vdash (\Gamma \vdash \Delta, \varphi)$.
8. *If* $d \Vdash (\Gamma, \exists R\varphi \vdash \Delta)$, *then there is* $d_\varphi \Vdash (\Gamma, \varphi \vdash \Delta)$.

*Furthermore, in all cases* $|d_\theta| \leq |d| + 1$ *and* $\rho_i(d_\theta) \leq \rho_i(d)$ *for* $i = 1, 2$ *and* $\theta = \varphi, \psi$.

*Proof.* The proof of all the results is similar, so we detail only (2). We proceed by structural induction on $d$; there are four different cases depending on the last rule applied in $d$.

- $d$ consists on an application of (Ax): if $\Gamma \cap \Delta$ contains a formula other than $\varphi \wedge \psi$, then both

  $$\dfrac{}{\Gamma \vdash \Delta, \varphi}\,(\mathrm{Ax}) \quad \text{and} \quad \dfrac{}{\Gamma \vdash \Delta, \psi}\,(\mathrm{Ax})$$

  are valid derivations satisfying the required properties.

  Else,

  $$\dfrac{\dfrac{}{\Gamma, \varphi \vdash \Delta, \varphi}\,(\mathrm{Ax})}{\Gamma \vdash \Delta, \varphi}\,(\wedge\ \mathrm{L}_1) \quad \text{and} \quad \dfrac{\dfrac{}{\Gamma, \psi \vdash \Delta, \psi}\,(\mathrm{Ax})}{\Gamma \vdash \Delta, \psi}\,(\wedge\ \mathrm{L}_2)$$

  are valid derivations again satisfying the required properties.

- The last rule applied in $d$ is ($\wedge$ R) with principal formula $\varphi \wedge \psi$: without loss of generality assume that $\varphi \wedge \psi$ is a side-formula of this rule, if necessary applying the Weakening Lemma 2. Then $d$ is of the form

  $$\dfrac{\dfrac{d_1}{\Gamma \vdash \Delta, \varphi \wedge \psi, \varphi} \qquad \dfrac{d_2}{\Gamma \vdash \Delta, \varphi \wedge \psi, \psi}}{\Gamma \vdash \Delta, \varphi \wedge \psi}\,(\wedge\ \mathrm{R})$$

  and applying the induction hypothesis to $d_1$ and $d_2$ yields the desired derivations.

- The last rule applied in $d$ is ($\wedge$ R) with principal formula other than $\varphi \wedge \psi$, or a different binary rule: then $d$ is of the form

  $$\dfrac{\dfrac{d_1}{\Gamma_1 \vdash \Delta_1, \varphi \wedge \psi} \qquad \dfrac{d_2}{\Gamma_2 \vdash \Delta_2, \varphi \wedge \psi}}{\Gamma \vdash \Delta, \varphi \wedge \psi}\,\mathrm{r}$$

  and by induction hypothesis there exist derivations $d_{1\varphi} \Vdash (\Gamma_1 \vdash \Delta_1, \varphi)$ and $d_{2\varphi} \Vdash (\Gamma_2 \vdash \Delta_2, \varphi)$, from which one can build $d_\varphi$ as

  $$\dfrac{\dfrac{d_{1\varphi}}{\Gamma_1 \vdash \Delta_1, \varphi} \qquad \dfrac{d_{2\varphi}}{\Gamma_2 \vdash \Delta_2, \varphi}}{\Gamma \vdash \Delta, \varphi}\,\mathrm{r}$$

and since $\left|d_{1\varphi}\right| \leq \left|d_{1\varphi}\right| + 1$ and $\rho_i\left(d_{1\varphi}\right) \leq \rho_i\left(d_{1\varphi}\right)$ for $i = 1,2$, and likewise for $d_{2\varphi}$, it follows that $d_\varphi$ fulfills the required conditions (in particular, this also holds if r is (Cut)). The reasoning for building $d_\psi$ is similar.

- The last rule applied in $d$ is a unary rule: then $d$ is of the form

$$\dfrac{\dfrac{d'}{\Gamma' \vdash \Delta', \varphi \wedge \psi}}{\Gamma \vdash \Delta, \varphi \wedge \psi}\,\mathrm{r}$$

and again by induction hypothesis there exists a derivation $d'_\varphi \Vdash (\Gamma' \vdash \Delta', \varphi)$ from which one can build $d_\varphi$ as

$$\dfrac{\dfrac{d'_\varphi}{\Gamma' \vdash \Delta', \varphi}}{\Gamma \vdash \Delta, \varphi}\,\mathrm{r}$$

and since $\left|d'_\varphi\right| \leq |d'| + 1$ and $\rho_i\left(d'_\varphi\right) \leq \rho_i(d')$ for $i = 1,2$, it follows that $d_\varphi$ fulfills the required conditions. The reasoning for building $d'_\psi$ is similar. $\square$

We now prove that any derivable sequent can be proved by means of a derivation of second-order rank 0.

**Lemma 6** (Second-order reduction lemma)**.** *Suppose that $d_1$ and $d_2$ are derivations with $d_1 \Vdash \Gamma_1 \vdash \varphi, \Delta_1$, $d_2 \Vdash \Gamma_2, \varphi \vdash \Delta_2$, $\rho_2(d_1) < |\varphi|_2$ and $\rho_2(d_2) < |\varphi|_2$. Then there exists a derivation d such that:*

- $d \Vdash (\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2)$*;*
- $|d| \leq 2(|d_1| + |d_2|)$*;*
- $\rho_2(d) < |\varphi|_2$.

*Proof.* Observe that $|\varphi|_2 > 0$, so $\varphi$ is not a first-order abstract. The proof is by induction on $|d_1| + |d_2|$.
(1) If $|d_1| + |d_2| = 0$: then both $d_1$ and $d_2$ are instances of (Ax). There are two possibilities.
(1a) If $\Gamma_1 \cap \Delta_1 \neq \emptyset$ or $\Gamma_2 \cap \Delta_2 \neq \emptyset$, then $d$ can be

$$\dfrac{}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\,(\mathrm{Ax})$$

which is a valid derivation trivially satisfying the thesis.
(1b) Otherwise, $\varphi \in \Gamma_1$ and $\varphi \in \Delta_2$, and the same derivation $d$ is valid.
(2) If $|d_1| + |d_2| > 0$, then there are three cases to consider.
(2a) $\varphi$ is not a principal formula in the last step of $d_1$. Then $d_1$ has the following form.

$$\dfrac{\dfrac{d'_{1i}}{\Gamma_1^i \vdash \Delta_1^i, \varphi}}{\Gamma_1 \vdash \Delta_1, \varphi}\,\mathrm{r}$$

(If r is unary, then $i = 1$, else $i$ can be 1 or 2.)

By induction hypothesis applied to $d'_{1i}$ and $d_2$, there exist derivations $d_i^*$ such that

- $d_i^* \Vdash \Gamma_1^i, \Gamma_2 \vdash \Delta_1^i, \Delta_2$;
- $|d_i^*| \leq 2(|d'_{1i}| + |d_2|)$;
- $\rho_2(d_i^*) < |\varphi|_2$.

Then we can take $d$ to be

$$\dfrac{\dfrac{d_i^*}{\Gamma_1^i, \Gamma_2 \vdash \Delta_1^i, \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\,\mathrm{r}$$

if necessary changing the variables in $d_i^*$ – in case r is ($\forall$ R), ($\exists$ L), ($\forall^2$ R) or ($\exists^2$ L) with an eigenvariable occuring free in $\Gamma_2 \cup \Delta_2$. Then

$$\begin{aligned} |d| &= \max(|d_i^*|) + 1 \\ &\leq 2\left(\max\left(\left|d'_{1i}\right|\right) + |d_2|\right) + 1 \\ &< 2(|d_1| + |d_2|) \end{aligned}$$

and, if r is not (Cut),

$$\rho_2(d) = \max\{\rho_2(d_i^*)\} < |\varphi|_2$$

or, if r is (Cut),

$$\rho_2(d) = \max(\{\rho_2(d_i^*)\} \cup \{|\theta|_2\})$$

where $|\theta|_2 < |\varphi|_2$, so again $\rho_2(d) < |\varphi|_2$.
(2b) $\varphi$ is not a principal formula in the last step of $d_2$: similar.
(2c) $\varphi$ is the principal formula in the last step of both $d_1$ and $d_2$. We have to look at the possible combinations of rules applied in the last step of $d_1$ and $d_2$. There are ten possible cases.
(2c.i) $d_1$ ends with ($\neg$ R) and $d_2$ with ($\neg$ L): then $\varphi$ is $\neg\psi$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_1$, eventually applying the Weakening Lemma 2. Then $d_1$ has the form

$$\dfrac{\dfrac{d'_1}{\Gamma_1, \psi \vdash \Delta_1, \varphi}}{\Gamma_1 \vdash \Delta_1, \varphi}\,\neg\,\mathrm{R}$$

Applying the induction hypothesis to $d'_1$ and $d_2$, we find a derivation $d' \Vdash (\Gamma_1, \Gamma_2, \psi \vdash \Delta_1, \Delta_2)$ such that

$$\left|d'\right| \leq 2\left(\left|d'_1\right| + |d_2|\right)$$
$$\rho_2\left(d'\right) < |\varphi|_2$$

Applying the Inversion Lemma 5 to $d_2$, we find a derivation $d_\psi \Vdash (\Gamma_2 \vdash \Delta_2, \psi)$ such that

$$\left|d_\psi\right| \leq |d_2| + 1$$
$$\rho_2\left(d_\psi\right) \leq \rho_2(d_2)$$

Take $d$ to be the following derivation.

$$\dfrac{\dfrac{d_\psi}{\Gamma_2 \vdash \Delta_2, \psi} \qquad \dfrac{d'}{\Gamma_1, \Gamma_2, \psi \vdash \Delta_1, \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\,(\mathrm{Cut})$$

Then:

$$\begin{aligned}|d| &= \max\left(\left|d_\psi\right|,\left|d'\right|\right)+1\\ &\le 2\left(|d_1|+|d_2|\right)\\ \rho_2(d) &= \max\left(\rho_2\left(d_\psi\right),\rho_2\left(d'\right),|\psi|_2\right)\\ &< |\varphi|_2\end{aligned}$$

(2c.ii) $d_1$ ends with ($\wedge$ R) and $d_2$ with ($\wedge$ L$_1$): then $\varphi$ is $\psi_1 \wedge \psi_2$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_2$, eventually applying the Weakening Lemma 2. Then $d_2$ has the form

$$\dfrac{\dfrac{d_2'}{\Gamma_2,\varphi,\psi_1 \vdash \Delta_2}}{\Gamma_2,\varphi \vdash \Delta_2}\ \wedge \mathrm{L}_1$$

Applying the induction hypothesis to $d_1$ and $d_2'$, we find a derivation $d' \Vdash (\Gamma_1,\Gamma_2,\psi_1 \vdash \Delta_1,\Delta_2)$ such that

$$\begin{aligned}\left|d'\right| &\le 2\left(|d_1|+\left|d_2'\right|\right)\\ \rho_2\left(d'\right) &< |\varphi|_2\end{aligned}$$

Applying the Inversion Lemma 5 to $d_1$, we find a derivation $d_{\psi_1} \Vdash (\Gamma_1 \vdash \Delta_1,\psi_1)$ such that

$$\begin{aligned}\left|d_{\psi_1}\right| &\le |d_1|+1\\ \rho_2\left(d_{\psi_1}\right) &\le \rho_2\left(d_1\right)\end{aligned}$$

Take $d$ to be the following derivation.

$$\dfrac{\dfrac{d_{\psi_1}}{\Gamma_1 \vdash \Delta_1,\psi_1} \qquad \dfrac{d'}{\Gamma_1,\Gamma_2,\psi_1 \vdash \Delta_1,\Delta_2}}{\Gamma_1,\Gamma_2 \vdash \Delta_1,\Delta_2}\ \text{(Cut)}$$

Then:

$$\begin{aligned}|d| &= \max\left(\left|d_{\psi_1}\right|,\left|d'\right|\right)+1\\ &\le 2\left(|d_1|+|d_2|\right)\\ \rho_2(d) &= \max\left(\rho_2\left(d_{\psi_1}\right),\rho_2\left(d'\right),|\psi_1|_2\right)\\ &< |\varphi|_2\end{aligned}$$

(2c.iii) $d_1$ ends with ($\wedge$ R) and $d_2$ with ($\wedge$ L$_2$): similar.
(2c.iv) $d_1$ ends with ($\vee$ R$_1$) and $d_2$ with ($\vee$ L): then $\varphi$ is $\psi_1 \vee \psi_2$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_1$, eventually applying the Weakening Lemma 2. Then $d_1$ has the form

$$\dfrac{\dfrac{d_1'}{\Gamma_1 \vdash \Delta_1,\varphi,\psi_1}}{\Gamma_1 \vdash \Delta_1,\varphi}\ \vee \mathrm{R}_1$$

Applying the induction hypothesis to $d_1'$ and $d_2$, we find a derivation $d' \Vdash (\Gamma_1,\Gamma_2 \vdash \Delta_1,\Delta_2,\psi_1)$ such that

$$\begin{aligned}\left|d'\right| &\le 2\left(\left|d_1'\right|+|d_2|\right)\\ \rho_2\left(d'\right) &< |\varphi|_2\end{aligned}$$

Applying the Inversion Lemma 5 to $d_2$, we find a derivation $d_{\psi_1} \Vdash (\Gamma_2,\psi_1 \vdash \Delta_2)$ such that

$$\begin{aligned}\left|d_{\psi_1}\right| &\le |d_2|+1\\ \rho_2\left(d_{\psi_1}\right) &\le \rho_2\left(d_2\right)\end{aligned}$$

Take $d$ to be the following derivation.

$$\dfrac{\dfrac{d'}{\Gamma_1,\Gamma_2 \vdash \Delta_1,\Delta_2,\psi_1} \qquad \dfrac{d_{\psi_1}}{\Gamma_2,\psi_1 \vdash \Delta_2}}{\Gamma_1,\Gamma_2 \vdash \Delta_1,\Delta_2}\ \text{(Cut)}$$

Then:

$$\begin{aligned}|d| &= \max\left(\left|d'\right|,\left|d_{\psi_1}\right|\right)+1\\ &\le 2\left(|d_1|+|d_2|\right)\\ \rho_2(d) &= \max\left(\rho_2\left(d'\right),\rho_2\left(d_{\psi_1}\right),|\psi_1|_2\right)\\ &< |\varphi|_2\end{aligned}$$

(2c.v) $d_1$ ends with ($\vee$ R$_2$) and $d_2$ with ($\vee$ L): similar.
(2c.vi) $d_1$ ends with ($\to$ R) and $d_2$ with ($\to$ L): then $\varphi$ is $\psi_1 \to \psi_2$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_1$, eventually applying the Weakening Lemma 2. Then $d_1$ has the form

$$\dfrac{\dfrac{d_1'}{\Gamma_1,\psi_1 \vdash \Delta_1,\varphi,\psi_2}}{\Gamma_1 \vdash \Delta_1,\varphi}\ (\to \mathrm{R})$$

Applying the induction hypothesis to $d_1'$ and $d_2$, we find a derivation $d' \Vdash (\Gamma_1,\Gamma_2,\psi_1 \vdash \Delta_1,\Delta_2,\psi_2)$ such that

$$\begin{aligned}\left|d'\right| &\le 2\left(\left|d_1'\right|+|d_2|\right)\\ \rho_2\left(d'\right) &< |\varphi|_2\end{aligned}$$

Applying the Inversion Lemma 5 to $d_2$, we find derivations $d_{\psi_1} \Vdash (\Gamma_2 \vdash \Delta_2,\psi_1)$ and $d_{\psi_2} \Vdash (\Gamma_2,\psi_2 \vdash \Delta_2)$ such that

$$\begin{aligned}\left|d_{\psi_i}\right| &\le |d_2|+1\\ \rho_2\left(d_{\psi_i}\right) &\le \rho_2\left(d_2\right)\end{aligned}$$

Take $d$ to be the following derivation.

$$\dfrac{\dfrac{d_{\psi_1}}{\Gamma_2 \vdash \Delta_2,\psi_1} \qquad \dfrac{\dfrac{d'}{\Gamma_1,\Gamma_2,\psi_1 \vdash \Delta_1,\Delta_2,\psi_2} \qquad \dfrac{d_{\psi_2}}{\Gamma_2,\psi_2 \vdash \Delta_2}}{\Gamma_1,\Gamma_2,\psi_1 \vdash \Delta_1,\Delta_2}\ \text{(Cut)}}{\Gamma_1,\Gamma_2 \vdash \Delta_1,\Delta_2}\ \text{(Cut)}$$

Then:

$$\begin{aligned}|d| &= \max\left(\left|d'\right|+2,\left|d_{\psi_1}\right|+1,\left|d_{\psi_2}\right|+2\right)\\ &\le \max\left(2\left(\left|d_1'\right|+|d_2|\right)+2,|d_2|+3\right)\\ &\le 2\left(|d_1|+|d_2|\right)\end{aligned}$$

noting that both $d_1$ and $d_2$ have length at least 1, and

$$\begin{aligned}\rho_2(d) &= \max\left(\rho_2\left(d'\right),\rho_2\left(d_{\psi_i}\right),|\psi_1|_2,|\psi_2|_2\right)\\ &< |\varphi|_2\end{aligned}$$

(2c.vii) $d_1$ ends with ($\forall$ R) and $d_2$ with ($\forall$ L): then $\varphi$ is $\forall x \psi$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_2$, eventually applying the Weakening Lemma 2. Then $d_2$ has the form

$$\dfrac{\dfrac{d_2'}{\Gamma_2, \varphi, \psi[s/x] \vdash \Delta_2}}{\Gamma_2, \varphi \vdash \Delta_2} \forall \mathrm{L}$$

Applying the induction hypothesis to $d_1$ and $d_2'$, we find a derivation $d' \Vdash (\Gamma_1, \Gamma_2, \psi[s/x] \vdash \Delta_1, \Delta_2)$ such that

$$\begin{aligned} |d'| &\leq 2\left(|d_1| + |d_2'|\right) \\ \rho_2\left(d'\right) &< |\varphi|_2 \end{aligned}$$

Applying the Inversion Lemma 5 to $d_1$, we find a derivation $d_\psi \Vdash (\Gamma_1 \vdash \Delta_1, \psi)$ such that

$$\begin{aligned} |d_\psi| &\leq |d_1| + 1 \\ \rho_2\left(d_\psi\right) &\leq \rho_2\left(d_1\right) \end{aligned}$$

By the First-order Substitution Lemma 3, the derivation $d_\psi[s/x]$ has the same length and second-order cut-rank as $d_\psi$, and furthermore $d_\psi[s/x] \Vdash (\Gamma_1 \vdash \Delta_1, \psi[s/x])$. Then take $d$ to be the following derivation.

$$\dfrac{\dfrac{d_\psi[s/x]}{\Gamma_1 \vdash \Delta_1, \psi[s/x]} \qquad \dfrac{d'}{\Gamma_1, \Gamma_2, \psi[s/x] \vdash \Delta_1, \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2} (\mathrm{Cut})$$

Then:

$$\begin{aligned} |d| &= \max\left(\left|d_\psi[s/x]\right|, \left|d'\right|\right) + 1 \\ &\leq 2\left(|d_1| + |d_2|\right) \\ \rho_2(d) &= \max\left(\rho_2\left(d_\psi[s/x]\right), \rho_2\left(d'\right), |\psi[s/x]|_2\right) \\ &< |\varphi|_2 \end{aligned}$$

noting that $|\psi[s/x]|_2 = |\psi|_2$.

(2c.viii) $d_1$ ends with ($\exists$ R) and $d_2$ with ($\exists$ L): then $\varphi$ is $\exists x \psi$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_1$, eventually applying the Weakening Lemma 2. Then $d_1$ has the form

$$\dfrac{\dfrac{d_1'}{\Gamma_1 \vdash \Delta_1, \varphi, \psi[s/x]}}{\Gamma_1, \varphi \vdash \Delta_1} \exists \mathrm{R}$$

Applying the induction hypothesis to $d_1'$ and $d_2$, we find a derivation $d' \Vdash (\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, \psi[s/x])$ such that

$$\begin{aligned} |d'| &\leq 2\left(|d_1'| + |d_2|\right) \\ \rho_2\left(d'\right) &< |\varphi|_2 \end{aligned}$$

Applying the Inversion Lemma 5 to $d_2$, we find a derivation $d_\psi \Vdash (\Gamma_2, \psi \vdash \Delta_2)$ such that

$$\begin{aligned} |d_\psi| &\leq |d_2| + 1 \\ \rho_2\left(d_\psi\right) &\leq \rho_2\left(d_2\right) \end{aligned}$$

By the First-order Substitution Lemma 3, the derivation $d_\psi[s/x]$ has the same length and second-order cut-rank as $d_\psi$, and furthermore $d_\psi[s/x] \Vdash (\Gamma_2, \psi[s/x] \vdash \Delta_2)$. Then take $d$ to be the following derivation.

$$\dfrac{\dfrac{d'}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, \psi[s/x]} \qquad \dfrac{d_\psi[s/x]}{\Gamma_2, \psi[s/x] \vdash \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2} (\mathrm{Cut})$$

Then:

$$\begin{aligned} |d| &= \max\left(\left|d'\right|, \left|d_\psi[s/x]\right|\right) + 1 \\ &\leq 2\left(|d_1| + |d_2|\right) \\ \rho_2(d) &= \max\left(\rho_2\left(d'\right), \rho_2\left(d_\psi[s/x]\right), |\psi[s/x]|_2\right) \\ &< |\varphi|_2 \end{aligned}$$

again noting that $|\psi[s/x]|_2 = |\psi|_2$.

(2c.ix) $d_1$ ends with ($\forall^2$ R) and $d_2$ with ($\forall^2$ L): then $\varphi$ is $\forall^2 R \psi$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_2$, eventually applying the Weakening Lemma 2. Then $d_2$ has the form

$$\dfrac{\dfrac{d_2'}{\Gamma_2, \varphi, \psi[\theta/R] \vdash \Delta_2}}{\Gamma_2, \varphi \vdash \Delta_2} \forall^2 \mathrm{L}$$

Applying the induction hypothesis to $d_1$ and $d_2'$, we find a derivation $d' \Vdash (\Gamma_1, \Gamma_2, \psi[\theta/R] \vdash \Delta_1, \Delta_2)$ such that

$$\begin{aligned} |d'| &\leq 2\left(|d_1| + |d_2'|\right) \\ \rho_2\left(d'\right) &< |\varphi|_2 \end{aligned}$$

Applying the Inversion Lemma 5 to $d_1$, we find a derivation $d_\psi \Vdash (\Gamma_1 \vdash \Delta_1, \psi)$ such that

$$\begin{aligned} |d_\psi| &\leq |d_1| + 1 \\ \rho_2\left(d_\psi\right) &\leq \rho_2\left(d_1\right) \end{aligned}$$

By the Second-order Substitution Lemma 4, the derivation $d_\psi[\theta/R]$ has the same length and second-order cut-rank as $d_\psi$, and furthermore $d_\psi[\theta/R] \Vdash (\Gamma_1 \vdash \Delta_1, \psi[\theta/R])$. Then take $d$ to be the following derivation.

$$\dfrac{\dfrac{d_\psi[\theta/R]}{\Gamma_1 \vdash \Delta_1, \psi[\theta/R]} \qquad \dfrac{d'}{\Gamma_1, \Gamma_2, \psi[\theta/R] \vdash \Delta_1, \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2} (\mathrm{Cut})$$

Then:

$$\begin{aligned} |d| &= \max\left(\left|d_\psi[\theta/R]\right|, \left|d'\right|\right) + 1 \\ &\leq 2\left(|d_1| + |d_2|\right) \\ \rho_2(d) &= \max\left(\rho_2\left(d_\psi[\theta/R]\right), \rho_2\left(d'\right), |\psi[\theta/R]|_2\right) \\ &< |\varphi|_2 \end{aligned}$$

using Lemma 1.

(2c.x) $d_1$ ends with ($\exists^2$ R) and $d_2$ with ($\exists^2$ L): then $\varphi$ is $\exists^2 R\psi$. Without loss of generality, assume that $\varphi$ is a side formula in the last step of $d_1$, eventually applying the Weakening Lemma 2. Then $d_1$ has the form

$$\dfrac{\dfrac{d_1'}{\Gamma_1 \vdash \Delta_1, \varphi, \psi[\theta/R]}}{\Gamma_1 \vdash \Delta_1, \varphi}\ \exists^2\text{ R}$$

Applying the induction hypothesis to $d_1'$ and $d_2$, we find a derivation $d' \Vdash (\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, \psi[\theta/R])$ such that

$$\begin{aligned} \left|d'\right| &\leq 2\left(\left|d_1'\right| + |d_2|\right) \\ \rho_2\left(d'\right) &< |\varphi|_2 \end{aligned}$$

Applying the Inversion Lemma 5 to $d_2$, we find a derivation $d_\psi \Vdash (\Gamma_2, \psi \vdash \Delta_2)$ such that

$$\begin{aligned} \left|d_\psi\right| &\leq |d_2| + 1 \\ \rho_2\left(d_\psi\right) &\leq \rho_2\left(d_2\right) \end{aligned}$$

By the Second-order Substitution Lemma 4, the derivation $d_\psi[\theta/R]$ has the same length and second-order cut-rank as $d_\psi$, and furthermore $d_\psi[\theta/R] \Vdash (\Gamma_2, \psi[\theta/R] \vdash \Delta_2)$. Then take $d$ to be the following derivation.

$$\dfrac{\dfrac{d'}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2, \psi[\theta/R]} \qquad \dfrac{d_\psi[\theta/R]}{\Gamma_2, \psi[\theta/R] \vdash \Delta_2}}{\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ (\text{Cut})$$

Then:

$$\begin{aligned} |d| &= \max\left(\left|d'\right|, \left|d_\psi[\theta/R]\right|\right) + 1 \\ &\leq 2\left(|d_1| + |d_2|\right) \\ \rho_2(d) &= \max\left(\rho_2\left(d'\right), \rho_2\left(d_\psi[\theta/R]\right), |\psi[\theta/R]|_2\right) \\ &< |\varphi|_2 \end{aligned}$$

again invoking Lemma 1. □

From this result it is now straightforward to prove that second-order formulas can be eliminated from cuts.

**Lemma 7** (Second-order cut-elimination 1)**.** *Suppose that $d \Vdash (\Gamma \vdash \Delta)$. If $\rho_2(d) > 0$, then there exists a derivation $d' \Vdash (\Gamma \vdash \Delta)$ such that $\rho_2\left(d'\right) < \rho_2(d)$ and $|d'| \leq 4^{|d|}$.*

*Proof.* By induction on $|d|$. If $|d| = 0$ then there is nothing to prove; otherwise, we proceed by case analysis on the last rule applied in $d$.

- If $d$ ends with the application of a unary rule, then $d$ is of the form

$$\dfrac{\dfrac{d'}{\Gamma' \vdash \Delta'}}{\Gamma \vdash \Delta}\ \text{r}$$

  with $\rho_2(d') = \rho_2(d)$. By induction hypothesis there is a derivation $d^* \vdash (\Gamma' \vdash \Delta')$ with $|d^*| \leq 4^{|d'|}$ and $\rho_2\left(d^*\right) < \rho_2(d)$. Then

$$\dfrac{\dfrac{d^*}{\Gamma' \vdash \Delta'}}{\Gamma \vdash \Delta}\ \text{r}$$

  is the required derivation.

- If $d$ ends with the application of a binary rule other than (Cut), then $d$ is of the form

$$\dfrac{\dfrac{d_1}{\Gamma_1 \vdash \Delta_1} \qquad \dfrac{d_2}{\Gamma_2 \vdash \Delta_2}}{\Gamma \vdash \Delta}\ \text{r}$$

  with $\rho_2\left(d_i\right) = \rho_2(d)$ for at least one of $i = 1, 2$. By induction hypothesis there are derivations $d_i' \vdash (\Gamma_i \vdash \Delta_i)$ with $|d_i'| \leq 4^{|d_i|}$ and $\rho_2\left(d_i'\right) < \rho_2(d)$ (if $\rho_2\left(d_i\right) < \rho_2(d)$, simply take $d_i' = d_i$). Then

$$\dfrac{\dfrac{d_1'}{\Gamma_1 \vdash \Delta_1} \qquad \dfrac{d_2'}{\Gamma_2 \vdash \Delta_2}}{\Gamma \vdash \Delta}\ \text{r}$$

  is the required derivation.

- If $d$ ends with the application of (Cut) with cut-formula $\varphi$ such that $|\varphi|_2 < \rho_2(d)$, then $d$ is of the form

$$\dfrac{\dfrac{d_1}{\Gamma_1 \vdash \Delta_1, \varphi} \qquad \dfrac{d_2}{\Gamma_2, \varphi \vdash \Delta_2}}{\Gamma \vdash \Delta}\ \text{r}$$

  with $\rho_2\left(d_i\right) = \rho_2(d)$ for at least one of $i = 1, 2$. By induction hypothesis there are derivations $d_i' \vdash (\Gamma_i \vdash \Delta_i)$ with $|d_i'| \leq 4^{|d_i|}$ and $\rho_2\left(d_i'\right) < \rho_2(d)$ (if $\rho_2\left(d_i\right) < \rho_2(d)$, simply take $d_i' = d_i$). Then

$$\dfrac{\dfrac{d_1'}{\Gamma_1 \vdash \Delta_1, \varphi} \qquad \dfrac{d_2'}{\Gamma_2, \varphi \vdash \Delta_2}}{\Gamma \vdash \Delta}\ \text{r}$$

  is the required derivation.

- If $d$ ends with the application of (Cut) with cut-formula $\varphi$ and $|\varphi|_2 = \rho_2(d)$, then $d$ is of the form

$$\dfrac{\dfrac{d_1}{\Gamma_1 \vdash \Delta_1, \varphi} \qquad \dfrac{d_2}{\Gamma_2, \varphi \vdash \Delta_2}}{\Gamma \vdash \Delta}\ \text{r}$$

  where possibly $\rho_2\left(d_i\right) = \rho_2(d)$ for at least one of $i = 1, 2$. If this is the case, then by induction hypothesis there are derivations $d_i' \vdash (\Gamma_i \vdash \Delta_i)$ with $|d_i'| \leq 4^{|d_i|}$ and $\rho_2\left(d_i'\right) < \rho_2(d)$; otherwise, simply take $d_i' = d_i$.

By the Reduction Lemma 6 there is a derivation $d' \Vdash (\Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2)$ such that

$$\begin{aligned}|d'| &\le 2\left(|d_1'| + |d_2'|\right)\\ &\le 2\left(4^{|d_1|} + 4^{|d_2|}\right)\\ &\le 2\left(4^{|d|-1} + 4^{|d|-1}\right) \le 4^{|d|}\end{aligned}$$

and $\rho_2(d') < |\varphi|_2 = \rho_2(d)$. □

**Theorem 1** (Second-order cut-elimination)**.** *Assume that $d \Vdash (\Gamma \vdash \Delta)$. Then there exists a derivation $d'$ of the same sequent such that $\rho_2(d') = 0$ and $|d'| \le 4^{|d|}_{\rho_2(d)}$, where $4^a_0 = 4^a$ and $4^a_{k+1} = 4^{4^a_k}$.*

*Proof.* By induction on $\rho_2(d)$. □

In order to obtain a cut-free proof, one simply repeats the process above, using $\rho_1$ instead of $\rho_2$ and restricting the three last results to derivations without second-order quantifiers in cut-formulas. This means that cases (2c.ix) and (2c.x) in the Reduction Lemma cannot occur; also, because of how first-order cut-rank is defined, the inequality $\rho_2(d) < |\varphi|_2$ is replaced by $\rho_1(d) \le |\varphi|_1$ (so $|\varphi|_1$ may be 0, unlike in the previous case).

We thus obtain the following result.

**Theorem 2** (First-order cut-elimination)**.** *Assume that $d \Vdash (\Gamma \vdash \Delta)$ and $\rho_2(d) = 0$. Then there exists a derivation $d'$ of the same sequent such that $\rho_1(d') = 0$ and $|d'| \le 4^{|d|}_{\rho_1(d)}$.*

Coupling Theorems 1 and 2 we achieve our goal.

**Theorem 3** (Cut-elimination)**.** *Assume that $d \Vdash (\Gamma \vdash \Delta)$. Then there exists a derivation $d'$ of the same sequent such that $\rho_1(d') = 0$ and $|d'| \le 4^{4^{|d|}_{\rho_2(d)}}_{\rho_1(d^*)}$, where $d^*$ is the derivation constructed from $d$ by applying Theorem 1.*

Explicit bounds for $\rho_1(d^*)$ in terms of $d$ can be obtained: since the cut-formulas of $d^*$ are obtained from the cut-formulas of $d$ by replacing second-order eigenvariables by their instances in applications of ($\forall^2$ L) or ($\exists^2$ R), their complexity is at most $k \times n^c$, where $k$ is the (regular) size of the most complex cut-formula in $d$, $n$ is the (first-order) size of the most complex first-order abstract in an instantiation of an eigenvariable of $d$, and $n$ is the maximum number of applications of ($\forall^2$ R) or ($\exists^2$ L) in a branch of $d$.

# 3 Final considerations

Formalizing these proofs in $\Delta I_0 + \mathsf{superexp}$ requires some care regarding the quantifiers in statements of the several lemmas. Interestingly, some of these raise issues even in the classical cut-elimination proof for first-order logic, which to the authors' knowledge have never been discussed.

First observe that, given a derivation $d$, one can recursively compute the (only) sequent it derives; so all hypotheses of the form "$d \Vdash \Gamma \vdash \Delta$" can be written down as recursive functions of $d$. Lemmas 2, 3 and 4 all pose no problem, since they directly define new derivations by a straightforward recursion and state recursive properties of this derivations. Similarly, Lemma 5 defines a new derivation by recursion using only decidable predicates for case analysis.

The only non-trivial case is that of Lemma 6, whose statement seemingly includes several nested quantifiers. In order to formulate and prove this statement in $\Delta I_0 + \mathsf{superexp}$, we need to bound the quantifiers over $d_1$, $d_2$ and $d$. Assuming some encoding on derivations, we can assume a fixed bound $k$ for the Gödel numbers of $d_1$ and $d_2$ (which also bounds $n$); then derivation $d$ has a maximum length (stated in the lemma), and furthermore all of its formulas have a maximum complexity (as discussed above), so the existential quantifier ranging over $d$ can be replaced by a bounded quantifier. Finally, one would perform an extra induction on $k$ to obtain the full result. The two cut-elimination lemmas can be treated similarly.

This analysis of the Reduction Lemma is completely independent of the second-order setting, and would already be required to argue that cut-elimination for first-order logic can be formalized in $\Delta I_0 + \mathsf{superexp}$.

Notice that the above results also hold if we allow for value-range terms in the syntax of the language (predicative or impredicative, it does not matter), as long as the calculus remains pure (without Law V), because the above proofs are independent of the structure of the terms.