HAL
open science

On the Proof Complexity of Deep Inference

Paola Bruscoli, Alessio Guglielmi

To cite this version:

Paola Bruscoli, Alessio Guglielmi. On the Proof Complexity of Deep Inference. ACM Transactions on Computational Logic, 2009, 10 (2), pp.34. 10.1145/1462179.1462186 . inria-00441211

HAL Id: inria-00441211

https://inria.hal.science/inria-00441211v1

Submitted on 15 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON THE PROOF COMPLEXITY OF DEEP INFERENCE

PAOLA BRUSCOLI AND ALESSIO GUGLIELMI

ABSTRACT. We obtain two results about the proof complexity of deep inference: 1) deep-inference proof systems are as powerful as Frege ones, even when both are extended with the Tseitin extension rule or with the substitution rule; 2) there are analytic deep-inference proof systems that exhibit an exponential speedup over analytic Gentzen proof systems that they polynomially simulate.

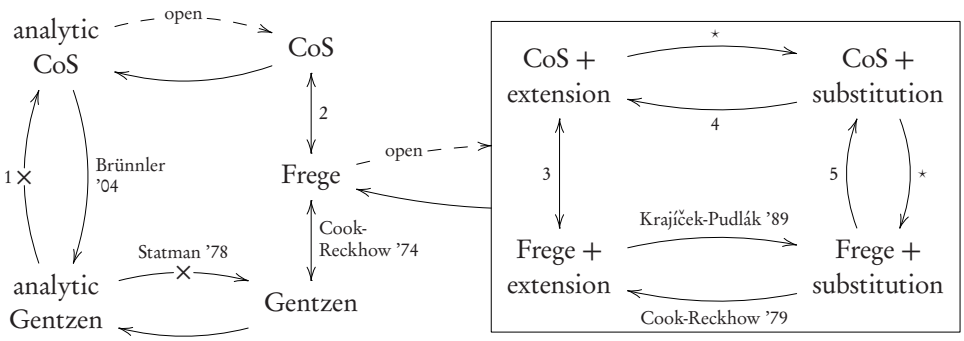
1. INTRODUCTION

Deep inference is a relatively new methodology in proof theory, consisting in dealing with proof systems whose inference rules are applicable at any depth inside formulae [Gug07b]. We obtain two results about the proof complexity of deep inference:

- deep-inference proof systems are as powerful as Frege ones, even when both are extended with the Tseitin extension rule or with the substitution rule;
- there are analytic deep-inference proof systems that exhibit an exponential speed-up over analytic Gentzen proof systems that they polynomially simulate.

These results are established for the *calculus of structures*, or *CoS*, the simplest formalism in deep inference [Gug07b], and in particular for its proof system SKS, introduced by Brünnler in [Brü04] and then extensively studied [Brü03a, Brü03b, Brü06a, Brü06d, BG04, BT01].

Our contributions fit in the following picture.



The notation $\mathscr{F} \longrightarrow \mathscr{G}$ indicates that formalism $\mathscr{F}$ polynomially simulates formalism $\mathscr{G}$; the notation $\mathscr{F} \not\longrightarrow \mathscr{G}$ indicates that it is known that this does not happen.

The left side of the picture represents, in part, the following. Analytic Gentzen systems, *i.e.*, Gentzen proof systems without the cut rule, can only prove certain formulae, which we call 'Statman tautologies', with proofs that grow exponentially in the size of the formulae. On the contrary, Gentzen systems with the cut rule can prove Statman tautologies by polynomially growing proofs. So, Gentzen systems p-simulate analytic

Date: April 19, 2009.

This research was partially supported by EPSRC grant EP/E042805/1 *Complexity and Non-determinism in Deep Inference*.

© ACM, 2009. This is the authors' version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in *ACM Transactions on Computational Logic* 10 (2:14) 2009, pp. 1–34, http://doi.acm.org/10.1145/1462179.1462186.

Gentzen ones, but not vice versa [Sta78]. Cook and Reckhow proved that Frege and Gentzen systems are polynomially equivalent, *i.e.*, each Frege system polynomially simulates any Gentzen system and vice versa [CR74].

In the box at the right of the figure, 'extension' refers to the Tseitin extension rule, and 'substitution' to the substitution rule. The works of Cook and Reckhow [CR79] and Krajíček and Pudlák [KP89] established that Frege + extension and Frege + substitution are polynomially equivalent. It is immediate to see that these formalisms polynomially simulate Frege and Gentzen, which, in turn, polynomially simulate analytic Gentzen. It is a major open problem to establish whether Frege polynomially simulates Frege + extension/substitution.

In this work, we establish the following results (numbered as in the previous figure):

(1) Analytic Gentzen does not polynomially simulate analytic CoS (essentially in the form of system SKS without cut); in fact, Statman tautologies admit polynomially growing proofs in analytic CoS (Theorem 3.12).
(2) CoS and Frege are polynomially equivalent (Theorems 4.7 and 4.12).
(3) There is a natural notion of (Tseitin) extension for CoS, and CoS + extension is polynomially equivalent to Frege + extension (Theorem 5.6).
(4) There is a natural notion of substitution for CoS, and CoS + substitution polynomially simulates CoS + extension (Theorem 5.12).
(5) Frege + substitution polynomially simulates CoS + substitution; this way, we know that all the extended formalisms are polynomially equivalent (Theorem 5.14).

The polynomial simulations indicated by $\star$ arcs in the picture follow from the others.

Establishing whether analytic CoS polynomially simulates CoS is an open problem.

After the necessary preliminaries, in Section 2, we see how CoS expresses Gentzen systems, including their properties, like analyticity, and then in Section 3 how it provides for exponentially more compact proofs. The relation between CoS and Frege systems is explored in Section 4 and the extensions are studied in Section 5. We conclude the article with a list of open problems, in Section 6.

2. Preliminaries

In this section, we quickly introduce the necessary deep-inference notions. A more extensive treatment of much of this material is in Brünnler's [Brü04].

We only need the following, basic proof complexity notions (see [CR79]).

Definition 2.1. A (*propositional*) *proof system* is a binary relation $\mathscr{S}$ between *formulae* α and *proofs* Π such that $\mathscr{S}$ is computable in polynomial time, and the formula α is a tautology if and only if there is a proof Π such that $\mathscr{S}(\alpha,\Pi)$; in this case we say that Π is a proof of α in $\mathscr{S}$. We say that proof system $\mathscr{S}$ *p-simulates* proof system $\mathscr{S}'$ if there is a polynomial-time computable algorithm that transforms every proof in $\mathscr{S}'$ into a proof in $\mathscr{S}$ of the same tautology. Two proof systems are *p-equivalent* if each p-simulates the other.

Remark 2.2. In the following, we state theorems on the existence of proofs in one proof system when proofs exist in another proof system, such that their size is polynomially related. Implicitly, we always mean that the new proofs are obtained by transforming the old ones by way of a polynomial-time computable algorithm.

Deep inference is a relatively recent development in proof theory. Its main idea is to provide a finer analysis of inference than possible with traditional methods, and one of the main objectives is to obtain a geometric semantics for proofs, inspired by linear logic's proof nets [Gir87]. Another objective is to provide a uniform and useful syntactic treatment of several logics, especially modal ones, for which no satisfactory proof theory existed before.

In deep inference, several formalisms can be defined with excellent structural properties, like locality for all the inference rules. The calculus of structures [Gug07b] is one of them and is now well developed for classical [Brü03a, Brü06a, Brü06d, BT01], intuitionistic [Tiu06a], linear [Str02, Str03b], modal [Brü06c, GT07, Sto07] and commutative/noncommutative logics [Gug07b, Tiu06b, Str03a, Bru02, DG04, GS01, GS02, GS09, Kah06b, Kah07]; for all these logics, quantification can be defined at any order. We emphasise that deep inference is developing the first reasonable proof theory for modal logics; the large number of different modal logic systems can be studied in simple and modular deep-inference systems, which are similar to their propositional logic counterparts and enjoy the same locality properties. The calculus of structures promoted the discovery of a new class of proof nets for classical and linear logic [LS05a, LS05b, LS06, SL04] (see also [Gui06]). Moreover, there exist implementations in Maude of deep-inference proof systems [Kah08].

In this article, we focus on the calculus of structures because it is well developed and is probably the simplest formalism definable in deep inference. The complexity results that we present here are not dependent on the choice of formalism; rather, they only depend on the deep-inference methodology and the finer granularity of inference rules that it yields. Adopting deep inference basically means that it is possible to replace subformulae inside formulae by other, implied subformulae, and that there is no limit to the nesting depth of subformulae. Formalisms like Gentzen's sequent calculus differ because they only rewrite formulae, or sequents, around their root connectives, and (we argue) they suffer excessive rigidity in the syntax and they do not sufficiently support geometric semantics.

Because of its geometric nature, it is important, in deep inference, to control whether propositional variables can be instantiated by formulae. In particular, normalisation (cut elimination) in deep inference crucially depends on the availability of 'atomic' inference rules, which are rules related to some topological invariants (see, for example, [GG08] for normalisation in propositional logic). In practice, we need two kinds of propositional variables: the atoms, only subject to renaming, and the formula variables, subject to (unrestricted) substitution. This distinction does not bear dramatic effects on proof complexity, but it does allow for some finer measures than otherwise possible.

Definition 2.3. *Formulae* of the *calculus of structures*, or *CoS*, are denoted by α, β, γ, δ and are freely built from: *units*, like f (false) and t (true); *atoms* a, b, c, d and $\bar{a}$, $\bar{b}$, $\bar{c}$, $\bar{d}$; (*formula*) *variables* A, B, C, D and $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{D}$; *logical relations*, like *disjunction* $[\alpha \vee \beta]$ and *conjunction* $(\alpha \wedge \beta)$. A formula is *ground* if it contains no variables. We usually omit external brackets of formulae, and sometimes we omit dispensable brackets under associativity. We use $\equiv$ to denote literal equality of formulae. The *size* $|\alpha|$ of a formula α is the number of unit, atom, and variable occurrences appearing in it. On the set of atoms, there is an involution $\bar{\cdot}$, called *negation* (*i.e.*, $\bar{\cdot}$ is a bijection from the set of atoms to itself such that $\bar{\bar{a}} \equiv a$); we require that $\bar{a} \not\equiv a$ for every a; when both a and $\bar{a}$ appear in a formula, we mean that atom a is mapped to $\bar{a}$ by $\bar{\cdot}$. An analogous involution is defined on the set of formula variables. The (*De Morgan*) *dual* of a formula is obtained by exchanging disjunction and conjunction and applying negation to all atoms and variables; we denote duals by using $\bar{\cdot}$; for example, the De Morgan dual of $\alpha \equiv \mathsf{t} \vee (a \wedge [\bar{B} \vee c])$ is $\bar{\alpha} \equiv \mathsf{f} \wedge [\bar{a} \vee (B \wedge \bar{c})]$. A *context* is a formula where one *hole* $\{\ \}$ appears in the place of a subformula; for example, $A \vee (b \wedge \{\ \})$ is a context; the generic context is denoted by $\xi\{\ \}$. The hole can be filled with formulae; for example, if $\xi\{\ \} \equiv b \wedge [\{\ \} \vee c]$, then $\xi\{a\} \equiv b \wedge [a \vee c]$, $\xi\{b\} \equiv b \wedge [b \vee c]$ and $\xi\{a \wedge B\} \equiv b \wedge [(a \wedge B) \vee c]$. The *size* of $\xi\{\ \}$ is defined as $|\xi\{\ \}| = |\xi\{a\}| - 1$.

Remark 2.4. We do not say that a is positive and $\bar{a}$ is negative. It only matters that, when a and $\bar{a}$ appear in the same formula, if one is negative the other is positive. In absence

of disambiguating information, there are two ways in which $\xi\{b\}$ might correspond to $b \wedge [b \vee c]$: one such that $\xi\{a\} \equiv a \wedge [b \vee c]$ and another such that $\xi\{a\} \equiv b \wedge [a \vee c]$.

The language of formulae is redundant because we can choose whether to use atoms or formula variables whenever a propositional variable is needed. The distinction between atoms and formula variables only plays a role in the choice of applicable inference rules, and this aspect is controlled by renamings and substitutions.

Definition 2.5. A *renaming* is a map from the set of atoms to itself, and is denoted by $\{a_1/b_1, a_2/b_2, \dots\}$; we use ρ for renamings. A renaming of α by $\rho = \{a_1/b_1, a_2/b_2, \dots\}$ is indicated by $\alpha\rho$ and is obtained by simultaneously substituting every occurrence of a_i in α by b_i and every occurrence of $\bar{a}_i$ by $\bar{b}_i$; for example, if $\alpha \equiv a \wedge [b \vee (a \wedge [\bar{a} \vee C])]$ then $\alpha\{a/\bar{b}, \bar{b}/c\} \equiv \bar{b} \wedge [\bar{c} \vee (\bar{b} \wedge [b \vee C])]$. A *substitution* is a map from the set of formula variables to formulae, denoted by $\{A_1/\beta_1, A_2/\beta_2, \dots\}$; we use σ for substitutions. An *instance* of α by $\sigma = \{A_1/\beta_1, A_2/\beta_2, \dots\}$ is indicated by $\alpha\sigma$ and is obtained by simultaneously substituting every occurrence of variable A_i in α by formula β_i and every occurrence of $\bar{A}_i$ by the De Morgan dual of β_i; for example, if $\alpha \equiv (A \wedge [A \vee \bar{A}]) \vee b$ then $\alpha\{A/(c \wedge \bar{B})\} \equiv ((c \wedge \bar{B}) \wedge [(c \wedge \bar{B}) \vee [\bar{c} \vee B]]) \vee b$.

Definition 2.6. A CoS (*inference*) *rule* ν is an expression $\nu\dfrac{\alpha}{\beta}$, where formulae α and β are called *premiss* and *conclusion*, respectively. A (*rule*) *instance* $\nu\dfrac{\gamma}{\delta}$ of $\nu\dfrac{\alpha}{\beta}$ is such that $\gamma \equiv \alpha\rho\sigma$ and $\delta \equiv \beta\rho\sigma$, for some renaming ρ and substitution σ. For some context $\xi\{\ \}$, a CoS (*inference*) *step*, *generated by* rule $\nu\dfrac{\alpha}{\beta}$ *via* its instance $\nu\dfrac{\gamma}{\delta}$, is the expression $\nu\dfrac{\xi\{\gamma\}}{\xi\{\delta\}}$.

Example 2.7. Given $\mathsf{i}{\downarrow}\dfrac{\mathsf{t}}{A \vee \bar{A}}$ and $\mathsf{ai}{\downarrow}\dfrac{\mathsf{t}}{a \vee \bar{a}}$, then $\mathsf{i}{\downarrow}\dfrac{\mathsf{t}}{(A \wedge b) \vee [\bar{A} \vee \bar{b}]}$ is an instance of $\mathsf{i}{\downarrow}$ and $\nu\dfrac{\mathsf{t}}{b \vee \bar{b}}$ is an instance of both $\mathsf{i}{\downarrow}$ and $\mathsf{ai}{\downarrow}$. The rule $\mathsf{c}{\downarrow}\dfrac{A \vee A}{A}$ generates the inference step $\mathsf{c}{\downarrow}\dfrac{[a \vee b] \wedge [(c \wedge D) \vee (c \wedge D)]}{[a \vee b] \wedge (c \wedge D)}$.

We usually classify deep-inference rules in three classes. For this, we rely on the notion of linearity, which, in this context, essentially means the same as in term-rewriting: a rewriting rule is linear if variables appear once in both sides of the rule. In other words, a linear rule does not create or destroy anything. These are the three classes of rules:

(1) *Atomic rules*. They usually correspond to structural rules in Gentzen systems; in normalisation and in semantics of proofs, they play a crucial role because they express the causality relations between atoms, so shaping the geometry of proofs. Their instances are obtained by renaming.
(2) *Noninvertible linear rules*. They usually correspond to logical rules in Gentzen systems. Since they are noninvertible, they express proper inference choices, but since they are linear, they do not alter the geometry of causality between atoms. Their instances are obtained by substitution.
(3) *Invertible linear rules*. These rules are equivalences between formulae that do not correspond to proper inference choices and have no impact on the geometry of proofs. For this reason, they are usually gathered into one big equivalence relation between formulae, corresponding to just one rule, defined via substitution.

The success of deep inference is due to its ability to separate rules into classes 1 and 2, which is only possible by adopting deep inference. The references to the 'geometry

Commutativity

$$\alpha \vee \beta = \beta \vee \alpha$$
$$\alpha \wedge \beta = \beta \wedge \alpha$$

Associativity

$$[\alpha \vee \beta] \vee \gamma = \alpha \vee [\beta \vee \gamma]$$
$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$$

Units

$$\alpha \vee \mathsf{f} = \alpha$$
$$\alpha \wedge \mathsf{t} = \alpha$$
$$\mathsf{t} \vee \mathsf{t} = \mathsf{t}$$
$$\mathsf{f} \wedge \mathsf{f} = \mathsf{f}$$

Context closure

if $\alpha = \beta$ then $\xi\{\alpha\} = \xi\{\beta\}$

FIGURE 1. Equality $=$ on formulae.

of proofs' can be understood by reading [GG08, LS05b]. Class 3 allows us to greatly simplify proofs and to hide, so to speak, a great deal of logical complexity (in the sense of size of proofs). We start by defining our 'class 3' rule, the others being dependent on specific proof systems.

Definition 2.8. The equality relation $=$ on formulae is defined by closing the equations in Figure 1 by reflexivity, symmetry, transitivity and by applying context closure. We define the inference rule $=$ as $=\dfrac{\alpha}{\beta}$, where $\alpha = \beta$.

The following remark helps in assessing how much complexity is hidden in $=$.

Remark 2.9. It is possible to decide $\alpha = \beta$ in polynomial time by reducing α and β to some canonical form and comparing the canonical forms. A canonical form under $=$ of any given formula can be obtained, for example, by removing as many units as possible and ordering units, atoms, and variables according to an arbitrary order; the canonical form is normal for associativity and units equations, when these are orientated from left to right. Let us assume a total order on the set of units, atoms, and variables; we now see in detail how to find an equivalent canonical formula in the case of a formula only containing one logical relation. On the formula, use commutativity until the minimal unit, atom, or variable appears in the leftmost position, then use associativity, orientated from left to right, until normality is reached. For example, on $[b \vee d] \vee [c \vee a]$, we perform the steps

$$[b \vee d] \vee [c \vee a] \rightsquigarrow [c \vee a] \vee [b \vee d] \rightsquigarrow [a \vee c] \vee [b \vee d] \rightsquigarrow a \vee [c \vee [b \vee d]] \quad .$$

This phase requires $\mathsf{O}(n)$ steps, where n is the size of the formula. We proceed the same way on the subformula immediately following the first element, and so on recursively; for example,

$$a \vee [c \vee [b \vee d]] \rightsquigarrow a \vee [[b \vee d] \vee c] \rightsquigarrow a \vee [b \vee [d \vee c]] \rightsquigarrow a \vee [b \vee [c \vee d]] \quad .$$

The number of steps of the algorithm for a formula only containing one logical relation is then $\mathsf{O}(n^2)$. On a generic formula, the same algorithm can be used, with the same number-of-steps complexity $\mathsf{O}(n^2)$ on the size n of the given formula, by adopting the lexicographic order induced by the given total order. This is an example, also involving

an initial $\mathsf{O}(n)$ phase of simplification of units:

$$
\begin{aligned}
&(\mathsf{t}\wedge\mathsf{t})\wedge[[a\vee([[b\vee d]\vee[c\vee a]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])]\vee\mathsf{f}]\\
\rightsquigarrow\ &\mathsf{t}\wedge[[a\vee([[b\vee d]\vee[c\vee a]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])]\vee\mathsf{f}]\\
\rightsquigarrow\ &[[a\vee([[b\vee d]\vee[c\vee a]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])]\vee\mathsf{f}]\wedge\mathsf{t}\\
\rightsquigarrow\ &[a\vee([[b\vee d]\vee[c\vee a]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])]\vee\mathsf{f}\\
\rightsquigarrow\ &a\vee([[b\vee d]\vee[c\vee a]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])\\
\rightsquigarrow^{\star}\ &a\vee([a\vee[b\vee[c\vee d]]]\wedge[[B\vee c]\vee[\mathsf{t}\vee a]])\\
\rightsquigarrow^{\star}\ &a\vee([a\vee[b\vee[c\vee d]]]\wedge[\mathsf{t}\vee[a\vee[c\vee B]]])\\
\rightsquigarrow\ &a\vee([\mathsf{t}\vee[a\vee[c\vee B]]]\wedge[a\vee[b\vee[c\vee d]]])\\
\rightsquigarrow\ &([\mathsf{t}\vee[a\vee[c\vee B]]]\wedge[a\vee[b\vee[c\vee d]]])\vee a \quad .
\end{aligned}
$$

This way we obtain a (unique, of course) canonical formula in $\mathsf{O}(n^2)$ steps, given any formula of size n, so we can decide the equivalence of two formulae α and β in $\mathsf{O}(n^2)$ steps, where $n=|\alpha|+|\beta|$. Notice that at each step the size of the formula stays the same or diminishes.

Definition 2.10. A CoS (*proof*) *system* is a finite set of inference rules. A CoS *derivation* Φ of *length* k in proof system $\mathscr{S}$, whose *premiss* is α_0 and *conclusion* is α_k, is a chain of inference steps

$$
\Phi=\begin{array}{c}\alpha_0\\ \hline \alpha_1\\ \hline \vdots\\ \hline \alpha_{k-1}\\ \hline \alpha_k\end{array}\ \begin{array}{l}\nu_1\\ \nu_2\\ \\ \nu_{k-1}\\ \nu_k\end{array} ,
$$

such that $\nu_1,\ldots,\nu_k$ is a sequence of inference rules that alternate between the $=$ rule and any rule of system $\mathscr{S}$, where $k\geq 0$. The same derivation can be indicated by $\begin{array}{c}\alpha_0\\ \Phi\|\mathscr{S}\\ \alpha_k\end{array}$, when the details are known or irrelevant; a *proof* is a derivation whose premiss is t. A derivation is *ground* if it contains no variables. Sometimes, we omit to indicate the inference steps generated by $=$. The *size* $|\Phi|$ of derivation Φ is the number of unit, atom, and variable occurrences appearing in it. We denote by $\xi\{\Phi\}$ the result of including every formula of Φ into the context $\xi\{\ \}$. We denote by $\Phi\rho$ and $\Phi\sigma$ the expression obtained from Φ by applying renaming ρ and substitution σ to every formula in Φ. A CoS proof system that, for every valid implication $\alpha\rightarrow\beta$, contains a derivation with premiss α and conclusion β, is said to be *implicationally complete*.

Remark 2.11. If Φ is a derivation, then $\xi\{\Phi\}$, $\Phi\rho$, and $\Phi\sigma$ are derivations, for every context $\xi\{\ \}$, renaming ρ, and substitution σ.

We use the notion of groundness to relate the complexity of deep-inference proof systems with atomic rules to proof systems without atomic rules, including those outside of deep inference. Due to the aforementioned redundancy in the language, groundness is not really a restriction.

Remark 2.12. Every nonground derivation can be transformed into an equivalent, ground one, by replacing variables with atoms in such a way that newly introduced atoms are different from the already present one.

Structural rules — *Logical* rule

$$\mathsf{SKSg}\left\{\begin{array}{llll}
\mathsf{i}\!\uparrow\dfrac{A\wedge\bar{A}}{\mathsf{f}} & \mathsf{w}\!\uparrow\dfrac{A}{\mathsf{t}} & \mathsf{c}\!\uparrow\dfrac{A}{A\wedge A} & \\
\textit{cointeraction} & \textit{coweakening} & \textit{cocontraction} & \\
\text{or } \textit{cut} & & & \\
\mathsf{i}\!\downarrow\dfrac{\mathsf{t}}{A\vee\bar{A}} & \mathsf{w}\!\downarrow\dfrac{\mathsf{f}}{A} & \mathsf{c}\!\downarrow\dfrac{A\vee A}{A} & \mathsf{s}\dfrac{A\wedge[B\vee C]}{(A\wedge B)\vee C} \\
\textit{interaction} & \textit{weakening} & \textit{contraction} & \textit{switch} \\
\text{or } \textit{identity} & & &
\end{array}\right\}\mathsf{KSg}$$

FIGURE 2. Systems SKSg and KSg.

Atomic structural rules — *Logical* rules

$$\mathsf{SKS}\left\{\begin{array}{lllll}
\mathsf{ai}\!\uparrow\dfrac{a\wedge\bar{a}}{\mathsf{f}} & \mathsf{aw}\!\uparrow\dfrac{a}{\mathsf{t}} & \mathsf{ac}\!\uparrow\dfrac{a}{a\wedge a} & & \\
\textit{cointeraction} & \textit{coweakening} & \textit{cocontraction} & & \\
\text{or } \textit{cut} & & & & \\
\mathsf{ai}\!\downarrow\dfrac{\mathsf{t}}{a\vee\bar{a}} & \mathsf{aw}\!\downarrow\dfrac{\mathsf{f}}{a} & \mathsf{ac}\!\downarrow\dfrac{a\vee a}{a} & \mathsf{s}\dfrac{A\wedge[B\vee C]}{(A\wedge B)\vee C} & \mathsf{m}\dfrac{(A\wedge B)\vee(C\wedge D)}{[A\vee C]\wedge[B\vee D]} \\
\textit{interaction} & \textit{weakening} & \textit{contraction} & \textit{switch} & \textit{medial} \\
\text{or } \textit{identity} & & & &
\end{array}\right\}\mathsf{KS}$$

FIGURE 3. Systems SKS and KS.

We can now define some deep-inference proof systems. System SKS is the most important for the proof theory of classical logic, because of its atomic structural rules. System SKSg relates SKS to proof systems in other formalisms, like Frege.

Definition 2.13. CoS proof systems $\mathsf{KSg} = \{\mathsf{i}\!\downarrow, \mathsf{w}\!\downarrow, \mathsf{c}\!\downarrow, \mathsf{s}\}$, $\mathsf{SKSg} = \mathsf{KSg} \cup \{\mathsf{i}\!\uparrow, \mathsf{w}\!\uparrow, \mathsf{c}\!\uparrow\}$, $\mathsf{KS} = \{\mathsf{ai}\!\downarrow, \mathsf{aw}\!\downarrow, \mathsf{ac}\!\downarrow, \mathsf{s}, \mathsf{m}\}$ and $\mathsf{SKS} = \mathsf{KS} \cup \{\mathsf{ai}\!\uparrow, \mathsf{aw}\!\uparrow, \mathsf{ac}\!\uparrow\}$ are defined in Figures 2 and 3, for a language containing f, t, disjunction, and conjunction. Proof systems where none of the rules $\mathsf{i}\!\uparrow$, $\mathsf{ai}\!\uparrow$, $\mathsf{w}\!\uparrow$, and $\mathsf{aw}\!\uparrow$ appear are said to be *analytic*.

Example 2.14. This is a valid derivation in all CoS proof systems defined previously (and it plays a role in the proof of Lemma 3.11):

$$\begin{array}{c}
=\dfrac{\gamma\vee[(([\bar{\alpha}\vee\alpha]\wedge c)\wedge(\alpha\wedge d))\vee\delta]}{\gamma\vee[(((\alpha\wedge d)\wedge c)\wedge[\alpha\vee\bar{\alpha}])\vee\delta]} \\
\mathsf{s}\dfrac{\gamma\vee[(((\alpha\wedge d)\wedge c)\wedge[\alpha\vee\bar{\alpha}])\vee\delta]}{\gamma\vee[[(((\alpha\wedge d)\wedge c)\wedge\alpha)\vee\bar{\alpha}]\vee\delta]} \\
=\dfrac{\gamma\vee[[(((\alpha\wedge d)\wedge c)\wedge\alpha)\vee\bar{\alpha}]\vee\delta]}{[\bar{\alpha}\vee\gamma]\vee[((\alpha\wedge c)\wedge(\alpha\wedge d))\vee\delta]}
\end{array}.$$

Note that SKSg, KSg, SKS, and KS are closed under renaming and substitution (see Remark 2.11). This is so because of the distinction between atoms and formula variables. Obtaining the closure of these and other systems under renaming and substitution is one of the main technical reasons for distinguishing between atoms and variables.

The following theorem is proved in [Brü04], and follows immediately from Section 3.1, where we prove that CoS systems p-simulate Gentzen systems.

Theorem 2.15. (Brünnler) *Systems* SKSg, KSg, SKS, *and* KS *are complete; systems* SKSg *and* SKS *are implicationally complete.*

The theorem holds also when restricting the language to ground derivations, since systems SKS and KS apply to them.

In the presence of cut, the coweakening and cocontraction rules do not play a major role in terms of proof complexity:

Theorem 2.16. *Systems* SKSg *and* KSg$\cup$i$\uparrow$ *are p-equivalent, and systems* SKS *and* KS$\cup$ai$\uparrow$ *are p-equivalent.*

Proof. Observe that the rules w$\uparrow$ and c$\uparrow$ can be derived in KSg$\cup$i$\uparrow$:

$$\dfrac{\dfrac{\dfrac{\dfrac{A}{A\wedge[\mathsf{f}\vee\mathsf{t}]}\,{=}}{(A\wedge\mathsf{f})\vee\mathsf{t}}\,\mathsf{s}}{(A\wedge\bar{A})\vee\mathsf{t}}\,\mathsf{w}{\downarrow}}{\mathsf{f}\vee\mathsf{t}}\,\mathsf{i}{\uparrow}}{\mathsf{t}}\,{=} \qquad \text{and} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A}{A\wedge\mathsf{t}}\,{=}}{A\wedge[[\bar{A}\vee\bar{A}]\vee(A\wedge A)]}\,\mathsf{i}{\downarrow}}{(A\wedge[\bar{A}\vee\bar{A}])\vee(A\wedge A)}\,\mathsf{s}}{(A\wedge\bar{A})\vee(A\wedge A)}\,\mathsf{c}{\downarrow}}{\mathsf{f}\vee(A\wedge A)}\,\mathsf{i}{\uparrow}}{A\wedge A}\,{=} \quad .$$

Similar constructions hold in KS$\cup$i$\uparrow$ for aw$\uparrow$ and ac$\uparrow$. □

It turns out that all the systems mentioned in the previous theorem are p-equivalent, as a consequence of Corollary 2.23.

Analytic systems are formally defined in Definition 2.13 for CoS, and 3.2 for Gentzen. Those definitions are specific to different systems in different formalisms, which is not necessarily satisfactory. Defining a general, syntax-independent concept of analyticity is a subject of ongoing research (see Problem 6.4). We briefly discuss now the notion of analyticity and its connections with the proof complexity of deep inference, as an introduction to our result on Statman tautologies.

A Gentzen system is said to be analytic when it does not contain the cut rule. Analytic Gentzen systems enjoy the 'subformula property', *i.e.*, proofs in these systems only contain subformulae of their conclusions. In fact, we might stipulate that enjoying the subformula property is a primitive notion of analyticity, which we can use to exclude the cut rule, as desired. In analytic Gentzen proofs, all formulae have lower or equal complexity than that of the conclusion, when complexity is measured, for example, as the and/or depth of a formula (*i.e.*, the number of alternations of conjunction and disjunction; see Definition 6.3). There is another property, of interest to us, that analytic Gentzen systems enjoy: given an inference rule and its conclusion, there are only finitely many premisses to choose from; we call such rules 'finitely generating'. The cut rule in Gentzen does not possess the subformula property nor is it finitely generating.

The primitive notion of analyticity that we are currently adopting for CoS is different from the one for Gentzen. We stipulate that a rule is analytic if its premiss is a formula obtained from a formula scheme by instantiating it with subformulae of the conclusion (so, the premiss is not just a subformula of the conclusion). This means that no atom or variable can appear in the premiss of an analytic rule that does not appear in its conclusion. It is, of course, a weaker condition than asking for the subformula property of Gentzen systems, but doing so is necessary if we want to adopt deep inference and obtain linear rules. Like the subformula property does for Gentzen, this weaker notion for CoS excludes the cut rule, but also the coweakening one. However, this is not an important difference with the sequent calculus because coweakening is irrelevant for the proof complexity of a CoS system (see, for example [GG08]). The reason for dealing with

coweakening is that, given the potential importance of cocontraction for proof complexity, we preferred to introduce top-down-symmetric CoS systems (so, closed by duality), even if coweakening and cocontraction are not required for completeness.

So, the two notions of analyticity, for Gentzen and for CoS, are such that the only important rules that are not analytic are the respective cut rules. Note that in both cases, analytic systems are made of finitely generating rules. However, there is an important difference: in CoS, the complexity of formulae in an analytic proof can be unboundedly greater than the complexity of the conclusion. Consider, for example, the derivation

$$\mathsf{c}{\downarrow}\,\dfrac{=\dfrac{\mathsf{s}\,\dfrac{c \vee (a \wedge [b \vee [c \vee (a \wedge b)])}{c \vee [(a \wedge b) \vee [c \vee (a \wedge b)]]}}{[c \vee (a \wedge b)] \vee [c \vee (a \wedge b)]}}{c \vee (a \wedge b)} \quad .$$

The and/or depth of the conclusion is 1, while that of the premiss is 3. We could repeat the construction on top of itself and further increase the and/or depth of the premiss at will.

Deep-inference systems can be top-down symmetric in the sense that a derivation can be flipped upside-down and negated and still be a valid derivation (we say that two such derivations are dual). Accordingly, some forms of analyticity can be defined in a symmetric way. Then, typically asymmetric theorems that depend on the notion of analyticity, like cut elimination, can be generalised to symmetric statements that imply cut elimination. This is not the place to be detailed about this aspect; suffice to say that we can obtain for CoS systems much stronger normalisation (and cut elimination) results than for Gentzen systems (see [Brü06b, GG08]).

As we see in Section 3.1, analyticity in CoS faithfully captures analyticity in Gentzen, in the sense that analytic CoS can produce isomorphic proofs to Gentzen ones (almost amounting to a change of notation). However, analytic CoS admits more proofs than analytic Gentzen, and, among CoS proofs, we can find some remarkably small ones, which analytic Gentzen cannot express; this is the subject of Section 3.2 on Statman tautologies.

Remark 2.17. The rules of SKS are *local*, in the sense that, for any language with a finite number of atoms, checking that a given expression is an instance of any of these rules requires time bounded by a constant (adopting a tree representation of formulae, for example). This property is peculiar to deep inference; it cannot be obtained in other formalisms. For example, a traditional, nonatomic contraction rule is not local because it requires checking the identity of two unbounded formulae. Contrary to other nonlocal rules, like identity in a Gentzen system, contraction cannot be replaced by its local, atomic counterpart without losing completeness. A counterexample showing this is in [Brü03b]. Locality can possibly lead to a new, general, productive notion of analyticity, as argued in Problem 6.4.

We conclude the section by showing the p-equivalence of systems with atomic rules to systems without atomic rules. We start by proving the result on ground derivations.

Lemma 2.18. *For every ground instance* $\mathsf{i}{\downarrow}\,\dfrac{\mathsf{t}}{\alpha \vee \bar{\alpha}}$ *there is a derivation* $\begin{array}{c}\mathsf{t}\\ \Phi\Big\Vert_{\{\mathsf{ai}\downarrow,\mathsf{s}\}}\\ \alpha \vee \bar{\alpha}\end{array}$ *and for every ground instance* $\mathsf{i}{\uparrow}\,\dfrac{\alpha \wedge \bar{\alpha}}{\mathsf{f}}$ *there is a derivation* $\begin{array}{c}\alpha \wedge \bar{\alpha}\\ \Phi\Big\Vert_{\{\mathsf{ai}\uparrow,\mathsf{s}\}}\\ \mathsf{f}\end{array}$*; in both cases* $|\Phi| \in \mathsf{O}(n^2)$*, where* $n = |\alpha|$.

Proof. Let us see the case for $\mathsf{i}{\uparrow}$, the other being its dual. We make an induction on the structure of α. The cases when α is a unit or an atom are trivial: in the former case Φ

consists of an instance of $=$ and in the latter the instance of $\mathsf{i}\!\uparrow$ is also an instance of $\mathsf{ai}\!\uparrow$. We only have to consider the case when $\alpha \equiv \beta \vee \gamma$: we apply the induction hypothesis on the derivation

$$\begin{array}{c} = \dfrac{[\beta \vee \gamma] \wedge (\bar{\beta} \wedge \bar{\gamma})}{} \\ \mathsf{s}\,\dfrac{(\bar{\beta} \wedge [\beta \vee \gamma]) \wedge \bar{\gamma}}{} \\ \mathsf{i}\!\uparrow \dfrac{[(\bar{\beta} \wedge \beta) \vee \gamma] \wedge \bar{\gamma}}{} \\ = \dfrac{[\mathsf{f} \vee \gamma] \wedge \bar{\gamma}}{} \\ \mathsf{i}\!\uparrow \dfrac{\gamma \wedge \bar{\gamma}}{\mathsf{f}} \end{array} ,$$

and we obtain a derivation whose length is $\mathsf{O}(n)$, and so its size is $\mathsf{O}(n^2)$, where $n = |\alpha|$. □

Lemma 2.19. *For every ground instance* $\mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\alpha}$ *there is a derivation* $\begin{array}{c}\mathsf{f}\\ \Phi \Big\| \{\mathsf{aw}\!\downarrow,\mathsf{s}\}\\ \alpha\end{array}$ *and for every ground instance* $\mathsf{w}\!\uparrow \dfrac{\alpha}{\mathsf{t}}$ *there is a derivation* $\begin{array}{c}\alpha\\ \Phi \Big\| \{\mathsf{aw}\!\uparrow,\mathsf{s}\}\\ \mathsf{t}\end{array}$*; in both cases* $|\Phi| \in \mathsf{O}(n^2)$*, where* $n = |\alpha|$.

Proof. The proof is similar to the one for Lemma 2.18. In case $\alpha \equiv \mathsf{t}$ an instance of $\mathsf{w}\!\downarrow$ yields

$$\begin{array}{c} = \dfrac{\mathsf{f}}{} \\ \mathsf{s}\,\dfrac{\mathsf{f} \wedge [\mathsf{f} \vee \mathsf{t}]}{} \\ = \dfrac{(\mathsf{f} \wedge \mathsf{f}) \vee \mathsf{t}}{\mathsf{t}} \end{array} ;$$

we can do similarly if $\alpha \equiv \mathsf{f}$ in an instance of $\mathsf{w}\!\uparrow$. These are the derivations for the inductive cases about $\mathsf{w}\!\downarrow$ (those about $\mathsf{w}\!\uparrow$ are dual):

$$\begin{array}{c} \mathsf{w}\!\downarrow \dfrac{\mathsf{f}}{\gamma} \\ = \dfrac{}{\mathsf{f} \vee \gamma} \\ \mathsf{w}\!\downarrow \dfrac{}{\beta \vee \gamma} \end{array} \quad \text{and} \quad \begin{array}{c} = \dfrac{\mathsf{f}}{} \\ \mathsf{w}\!\downarrow \dfrac{}{\mathsf{f} \wedge \mathsf{f}} \\ \mathsf{w}\!\downarrow \dfrac{}{\mathsf{f} \wedge \gamma} \\ \mathsf{w}\!\downarrow \dfrac{}{\beta \wedge \gamma} \end{array} .$$

From these we obtain a derivation whose length is $\mathsf{O}(n)$, and so its size is $\mathsf{O}(n^2)$, where $n = |\alpha|$. □

Remark 2.20. In the statement of Lemma 2.19, instead of $\{\mathsf{aw}\!\downarrow,\mathsf{s}\}$ and $\{\mathsf{aw}\!\uparrow,\mathsf{s}\}$ we could have used $\{\mathsf{aw}\!\downarrow,\mathsf{m}\}$ and $\{\mathsf{aw}\!\uparrow,\mathsf{m}\}$, respectively.

Lemma 2.21. *For every ground instance* $\mathsf{c}\!\downarrow \dfrac{\alpha \vee \alpha}{\alpha}$ *there is a derivation* $\begin{array}{c}\alpha \vee \alpha\\ \Phi \Big\| \{\mathsf{ac}\!\downarrow,\mathsf{m}\}\\ \alpha\end{array}$ *and for every ground instance* $\mathsf{c}\!\uparrow \dfrac{\alpha}{\alpha \wedge \alpha}$ *there is a derivation* $\begin{array}{c}\alpha\\ \Phi \Big\| \{\mathsf{ac}\!\uparrow,\mathsf{m}\}\\ \alpha \wedge \alpha\end{array}$*; in both cases* $|\Phi| \in \mathsf{O}(n^2)$*, where* $n = |\alpha|$.

Proof. The proof is similar to the one for Lemma 2.18. These are the derivations for the inductive cases about $\mathsf{c}\!\downarrow$ (those about $\mathsf{c}\!\uparrow$ are dual):

$$\begin{array}{c} = \dfrac{[\beta \vee \gamma] \vee [\beta \vee \gamma]}{} \\ \mathsf{c}\!\downarrow \dfrac{[\beta \vee \beta] \vee [\gamma \vee \gamma]}{} \\ \mathsf{c}\!\downarrow \dfrac{\beta \vee [\gamma \vee \gamma]}{\beta \vee \gamma} \end{array} \quad \text{and} \quad \begin{array}{c} \mathsf{m}\,\dfrac{(\beta \wedge \gamma) \vee (\beta \wedge \gamma)}{} \\ \mathsf{c}\!\downarrow \dfrac{[\beta \vee \beta] \wedge [\gamma \vee \gamma]}{} \\ \mathsf{c}\!\downarrow \dfrac{\beta \wedge [\gamma \vee \gamma]}{\beta \wedge \gamma} \end{array} .$$

From these we obtain a derivation whose length is $\mathsf{O}(n)$, and so its size is $\mathsf{O}(n^2)$, where $n = |\alpha|$. □

$$\mathsf{cut}\,\frac{\phi, A \quad \bar{A}, \psi}{\phi, \psi}\ \textit{cut}$$

$$\mathsf{id}\,\frac{}{A, \bar{A}}\ \textit{identity} \qquad \mathsf{t}\,\frac{}{\mathsf{t}}\ \textit{true} \qquad \mathsf{w}\,\frac{\phi}{\phi, A}\ \textit{weakening} \qquad \mathsf{c}\,\frac{\phi, A, A}{\phi, A}\ \textit{contraction} \qquad \vee\,\frac{\phi, A, B}{\phi, A \vee B}\ \textit{disjunction} \qquad \wedge\,\frac{\phi, A \quad B, \psi}{\phi, A \wedge B, \psi}\ \textit{conjunction}$$

FIGURE 4. System Gentzen.

Theorem 2.22. *For every ground* SKSg *derivation* Φ *there is a ground* SKS *derivation* Φ' *with the same premiss and conclusion of* Φ*; if* n *is the size of* Φ *then the size of* Φ' *is* $\mathsf{O}(n^2)$*; moreover, if* Φ *is in* KSg *then* Φ' *is in* KS.

Proof. The theorem follows immediately from Lemmas 2.18, 2.19, and 2.21. □

By Remark 2.12, every derivation can be 'grounded', so:

Corollary 2.23. KS *and* KSg *are p-equivalent and* SKS *and* SKSg *are p-equivalent.*

Remark 2.24. Sometimes, we use nonatomic structural rule instances in SKS and KS derivations: those instances actually stand for the SKS and KS derivations that would be obtained according to the proofs of Lemmas 2.18, 2.19, and 2.21. In this sense, we say that i↓, i↑, w↓, w↑, c↓, and c↑ are 'macro' rules for SKS and KS. The reason we might want to work with macro rules in SKS and KS instead of working in SKSg and KSg and then appealing to Theorem 2.22 is to obtain finer upper bounds. This is because the size of formulae over which nonatomic structural rules operate can be much smaller than the square root of the size of a derivation, which is the pessimistic assumption of Theorem 2.22.

Remark 2.25. All implicationally complete CoS proof systems are p-equivalent. This can be proved analogously to, or resorting to, a similar 'robustness' result for Frege systems (Theorem 4.2), as argued in Remark 4.13. This means that studying proof complexity for SKSg and SKS has universal value for all CoS systems for propositional logic.

3. CALCULUS OF STRUCTURES, GENTZEN PROOF SYSTEMS AND STATMAN TAUTOLOGIES

There are two parts in this section. In the first part, we show how CoS naturally p-simulates Gentzen systems, and in particular how it realizes Gentzen's notion of analyticity. In the second part, we show that analytic CoS admits polynomial proofs when analytic Gentzen only has exponential ones, in the case of Statman tautologies.

3.1. Calculus of Structures and Gentzen Proof Systems.

In this section, we adopt a specific one-sided (Gentzen-Schütte) sequent system that we call Gentzen (and that is called GS1p in [TS96]). We could have adopted any other style of presentation without affecting our results. In fact, for Gentzen systems an analogous 'robustness' theorem to that for Frege systems (Theorem 4.2) can be established. This means that studying the proof complexity of Gentzen has universal value for the class of Gentzen systems.

Definition 3.1. Over the language of SKS formulae, the sequent-calculus proof system Gentzen is defined by the *inference rules* in Figure 4, where ϕ and ψ stand for multisets of formulae and the symbol ',' represents multiset union. We interpret multisets of formulae as their disjunction (where associativity is irrelevant). *Derivations*, denoted by Δ, are trees obtained by composing instances of inference rules; the leaves of a derivation are

its *premisses* and the root is its *conclusion*; a derivation Δ with premisses $\phi_1, \ldots, \phi_h$ and conclusion ψ is denoted by

$$\begin{array}{c} \phi_1 \cdots \phi_h \\ \hline \Delta \\ \hline \psi \end{array} .$$

A derivation with no premisses is a *proof*. The *size* $|\Delta|$ of derivation Δ is the number of unit, atom, and variable occurrences appearing in it. In the following, every SKS formula is translated into a Gentzen formula in the obvious way, and vice versa; in particular, we translate a Gentzen multiset $\phi = \alpha_1, \ldots, \alpha_h$ into $\alpha_1 \vee \cdots \vee \alpha_h$.

In the language, we keep the distinction between atoms and variables because, thanks to atoms, we obtain a better upper bound for the size of Statman tautologies proofs, in the Section 3.2. As we said in the case of CoS, the redundancy in the language has no consequences outside of the possibility of using certain CoS rules instead of others.

Definition 3.2. The proof system *analytic* Gentzen is proof system Gentzen without the cut rule; *analytic* derivations and proofs are those derivations and proofs in Gentzen where no instances of the cut rule appear.

We know, of course, that both Gentzen and analytic Gentzen are complete, and that Gentzen proofs can be transformed into analytic Gentzen proofs by a cut-elimination procedure, which, in general, blows-up a given proof exponentially.

Every Gentzen derivation has natural counterparts in CoS: the idea is to (arbitrarily) sequentialise its tree structure. This is possible because the natural logical relation between tree branches is conjunction, which CoS can represent, of course. In doing so, we pay in terms of complexity because the tree structure is less redundant than CoS contexts: the size of derivations grows quadratically. Other deep-inference formalisms (currently under development, see [BL05, Gug04, Gug05]) are more efficient than CoS and Gentzen formalisms in dealing with this so-called 'bureaucracy'.

Remark 3.3. In the following, we assume that an empty conjunction can be represented by a nonempty conjunction of t units.

Theorem 3.4. *For every* Gentzen *derivation* Δ *with premisses* $\phi_1, \ldots, \phi_h$ *and conclusion* ψ *there is a derivation* $\begin{array}{c} \phi_1 \wedge \cdots \wedge \phi_h \\ \Phi \Big\| \mathsf{SKSg} \\ \psi \end{array}$ *; if* n *is the size of* Δ*, the size of* Φ *is* $\mathsf{O}(n^2)$*; moreover, if* Δ *is analytic then* Φ *is in* KSg.

Proof. We proceed by induction on the tree structure of Δ. The base cases $\mathsf{id}\,\dfrac{}{A, \bar{A}}$ and $\mathsf{t}\,\dfrac{}{\mathsf{t}}$ are, respectively, translated into $\mathsf{i}{\downarrow}\,\dfrac{\mathsf{t}}{A, \bar{A}}$ and t. The derivations

$$\begin{array}{c} \phi_1 \cdots \phi_h \\ \hline \Delta_1 \\ \hline \phi \\ \hline \mathsf{w}\,\phi, A \end{array} , \quad \begin{array}{c} \phi_1 \cdots \phi_h \\ \hline \Delta_1 \\ \hline \phi, A, A \\ \hline \mathsf{c}\,\phi, A \end{array} , \quad \text{and} \quad \begin{array}{c} \phi_1 \cdots \phi_h \\ \hline \Delta_1 \\ \hline \phi, A, B \\ \hline \vee\,\phi, A \vee B \end{array}$$

are, respectively, translated into

$$\begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_h \\ {\scriptstyle \Phi_1}\Big\|{\scriptstyle \mathsf{SKSg}} \\ \phi \\ \hline \phi \vee \mathsf{f} \\ \hline \phi \vee A\end{array}{\scriptstyle =}{\scriptstyle \mathsf{w}\downarrow} \quad , \quad \begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_h \\ {\scriptstyle \Phi_1}\Big\|{\scriptstyle \mathsf{SKSg}} \\ \phi \vee [A \vee A] \\ \hline \phi \vee A\end{array}{\scriptstyle \mathsf{c}\downarrow} \quad , \quad \text{and} \quad \begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_h \\ {\scriptstyle \Phi_1}\Big\|{\scriptstyle \mathsf{SKSg}} \\ \phi \vee [A \vee B]\end{array} \quad ,$$

where Φ_1 is obtained by induction from Δ_1, and some possibly necessary instances of the $=$ rule have been omitted (they depend on the exact translation of Gentzen multisets into SKSg formulae). The derivations

$$\begin{array}{c}\phi_1 \cdots \phi_h \quad \phi_{h+1} \cdots \phi_k \\ \Delta_1 \quad \Delta_2 \\ \phi, A \qquad B, \psi \\ \hline \phi, A \wedge B, \psi\end{array}{\scriptstyle \wedge} \quad \text{and} \quad \begin{array}{c}\phi_1 \cdots \phi_h \quad \phi_{h+1} \cdots \phi_k \\ \Delta_1 \quad \Delta_2 \\ \phi, A \qquad \bar{A}, \psi \\ \hline \phi, \psi\end{array}{\scriptstyle \mathsf{cut}}$$

are, respectively, translated into

$$\begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_k \\ {\scriptstyle \Phi_1 \wedge \Phi_2}\Big\| \\ [\phi \vee A] \wedge [B \vee \psi] \\ \hline [B \vee \psi] \wedge [A \vee \phi] \\ \hline ([B \vee \psi] \wedge A) \vee \phi \\ \hline \phi \vee (A \wedge [B \vee \psi]) \\ \hline \phi \vee [(A \wedge B) \vee \psi]\end{array}{\scriptstyle =}{\scriptstyle \mathsf{s}}{\scriptstyle =}{\scriptstyle \mathsf{s}} \quad \text{and} \quad \begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_k \\ {\scriptstyle \Phi_1 \wedge \Phi_2}\Big\| \\ [\phi \vee A] \wedge [\bar{A} \vee \psi] \\ \hline [\bar{A} \vee \psi] \wedge [A \vee \phi] \\ \hline ([\bar{A} \vee \psi] \wedge A) \vee \phi \\ \hline \phi \vee (A \wedge [\bar{A} \vee \psi]) \\ \hline \phi \vee [(A \wedge \bar{A}) \vee \psi] \\ \hline \phi \vee [\mathsf{f} \vee \psi] \\ \hline \phi \vee \psi\end{array}{\scriptstyle =}{\scriptstyle \mathsf{s}}{\scriptstyle =}{\scriptstyle \mathsf{s}}{\scriptstyle \mathsf{i}\uparrow}{\scriptstyle =} \quad ,$$

where Φ_1 and Φ_2 are obtained by induction from Δ_1 and Δ_2, some possibly necessary instances of the $=$ rule have been omitted, and $\Phi_1 \wedge \Phi_2$ stands for the derivation

$$\begin{array}{c}\phi_1 \wedge \cdots \wedge \phi_k \\ {\scriptstyle \Phi_1 \wedge (\phi_{h+1} \wedge \cdots \wedge \phi_k)}\Big\| \\ [\phi \vee A] \wedge (\phi_{h+1} \wedge \cdots \wedge \phi_k) \\ {\scriptstyle [\phi \vee A] \wedge \Phi_2}\Big\| \\ [\phi \vee A] \wedge [B \vee \psi]\end{array} \quad ,$$

where B is possibly instantiated by $\bar{A}$; the length of this derivation and the size of the largest formula appearing in it are both $\mathsf{O}(n)$. The resulting $\mathsf{O}(n^2)$ measure of these last two cases dominates the others. □

Corollary 3.5. SKSg *p-simulates* Gentzen *and* KSg *p-simulates analytic* Gentzen.

Although it does not explicitly address complexity, [Brü04] is more exhaustive than the aforesaid on the two-way translation between SKSg and Gentzen. Translating from SKSg to Gentzen crucially employs the cut rule: for every inference step in SKSg, a cut instance is used in Gentzen. So, while it is very natural and easy to show that Gentzen p-simulates SKSg (see [Brü04]), we are left with the question: does analytic Gentzen p-simulate KSg?

3.2. **Analytic Calculus of Structures on Statman Tautologies.** We prove here that analytic Gentzen does not p-simulate KSg. In fact, the CoS (polynomial) inefficiency in dealing with context bureaucracy is compensated by its freedom in applying inference rules, which leads to exponential speedups on certain classes of tautologies. Here, we study Statman tautologies, which have been used to provide the classic lower bound for analytic Gentzen systems: no cut-free proofs of Statman tautologies are possible in analytic Gentzen without the proofs growing exponentially over the size of the tautologies they prove [Sta78]. We show that, on the contrary, KSg and KS prove Statman tautologies with polynomially growing analytic proofs.

Remark 3.6. The subset of SKSg only containing analytic rules is equal to KSg plus the cocontraction rule. We do not know whether cocontraction provides for exponential speedups, so separating, proof-complexity-wise, the class of KSg from that of 'analytic CoS'; about this, see Problem 6.2. In our opinion, the very notion of analyticity would benefit from some further analysis; about this, see Problem 6.4.

Definition 3.7. For $n \geq 1$, consider the following formulae:

$$\begin{array}{ll} \alpha_i \equiv \bar{c}_i \vee \bar{d}_i & \text{for } i \geq 1, \\ \beta^n_k \equiv \bigwedge_{i=n}^{k} \alpha_i \equiv \alpha_n \wedge \beta^{n-1}_k & \text{for } n \geq k > 1, \\ \gamma^n_k \equiv \beta^n_{k+1} \wedge c_k & \text{for } n > k \geq 1, \\ \delta^n_k \equiv \beta^n_{k+1} \wedge d_k & \text{for } n > k \geq 1. \end{array}$$

Statman tautologies are, for $n \geq 1$, the formulae:

$$\mathsf{S}_n \equiv \bar{\alpha}_n \vee [(\gamma^n_{n-1} \wedge \delta^n_{n-1}) \vee [\cdots \vee [(\gamma^n_1 \wedge \delta^n_1) \vee \alpha_1] \cdots]] \quad .$$

Example 3.8. These are the first three Statman tautologies:

$$\begin{aligned} \mathsf{S}_1 &\equiv (c_1 \wedge d_1) \vee [\bar{c}_1 \vee \bar{d}_1] \quad , \\ \mathsf{S}_2 &\equiv (c_2 \wedge d_2) \vee [(([\bar{c}_2 \vee \bar{d}_2] \wedge c_1) \wedge ([\bar{c}_2 \vee \bar{d}_2] \wedge d_1)) \vee [\bar{c}_1 \vee \bar{d}_1]] \quad , \\ \mathsf{S}_3 &\equiv (c_3 \wedge d_3) \vee [(([\bar{c}_3 \vee \bar{d}_3] \wedge c_2) \wedge ([\bar{c}_3 \vee \bar{d}_3] \wedge d_2)) \vee \\ &\qquad [((([\bar{c}_3 \vee \bar{d}_3] \wedge [\bar{c}_2 \vee \bar{d}_2]) \wedge c_1) \wedge (([\bar{c}_3 \vee \bar{d}_3] \wedge [\bar{c}_2 \vee \bar{d}_2]) \wedge d_1)) \vee \\ &\qquad [\bar{c}_1 \vee \bar{d}_1]]] \quad . \end{aligned}$$

It is perhaps easier to understand their meaning by using implication, as in

$$\begin{aligned} \mathsf{S}'_3 \equiv [\bar{c}_3 \vee \bar{d}_3] \to ([([\bar{c}_3 \vee \bar{d}_3] \to \bar{c}_2) \vee ([\bar{c}_3 \vee \bar{d}_3] \to \bar{d}_2)] \to \\ ([(([\bar{c}_3 \vee \bar{d}_3] \wedge [\bar{c}_2 \vee \bar{d}_2]) \to \bar{c}_1) \vee (([\bar{c}_3 \vee \bar{d}_3] \wedge [\bar{c}_2 \vee \bar{d}_2]) \to \bar{d}_1)] \to \\ [\bar{c}_1 \vee \bar{d}_1])) \quad . \end{aligned}$$

Remark 3.9. $|\mathsf{S}_n| = 2 + 2\sum_{k=1}^{n-1} |\gamma^n_k| + 2 = 2\sum_{k=2}^{n} (|\beta^n_k| + 1) + 4 = 2n^2 + 2.$

It is not difficult to see why analytic Gentzen proofs of Statman tautologies grow exponentially (this is a classic argument that can be found in many textbooks; see, for example, [CK02]). Basically, all what analytic Gentzen can do while building a proof of S_n is to generate a proof tree with $\mathsf{O}(2^n)$ branches. The next lemma shows the crucial advantage of deep inference over Gentzen systems: Statman tautologies can be proved 'from the inside out', which is precisely what Gentzen systems can only do by resorting to convoluted proofs involving cuts (so, nonanalytic proofs).

Remark 3.10. In the following, for brevity, we label inference steps with expressions like $n \cdot \nu$, to denote n inference steps involving rule ν.

Lemma 3.11. *For Statman tautologies* S_n *and* S_{n+1} *there exists a derivation* $\begin{array}{c}\mathsf{S}_n\\ \Vert\mathsf{KS}\\ \mathsf{S}_{n+1}\end{array}$ *whose length is* $\mathsf{O}(n)$ *and size is* $\mathsf{O}(n^3)$.

Proof. We refer to Definition 3.7. The requested derivation is

$$\Phi = \dfrac{\dfrac{\dfrac{\begin{array}{l}(c_n \wedge d_n)\vee\\ [((\beta^n_n \wedge c_{n-1})\wedge(\beta^n_n \wedge d_{n-1}))\vee[\cdots\vee\\ [((\beta^n_2 \wedge c_1)\wedge(\beta^n_2 \wedge d_1))\vee\alpha_1]\cdots]]\end{array}}{\begin{array}{l}(([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge c_n)\wedge([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge d_n))\vee\\ [((([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge\beta^n_n)\wedge c_{n-1})\wedge(([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge\beta^n_n)\wedge d_{n-1}))\vee[\cdots\vee\\ [((([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge\beta^n_2)\wedge c_1)\wedge(([\alpha_{n+1}\vee\bar{\alpha}_{n+1}]\wedge\beta^n_2)\wedge d_1))\vee\alpha_1]\cdots]]\end{array}}\,2n\cdot\mathsf{i}{\downarrow}}{\begin{array}{l}\overbrace{[\bar{\alpha}_{n+1}\vee[\cdots\vee[\bar{\alpha}_{n+1}\vee\bar{\alpha}_{n+1}]\cdots]]}^{2n}\vee\\ [((\alpha_{n+1}\wedge c_n)\wedge(\alpha_{n+1}\wedge d_n))\vee\\ [((((\alpha_{n+1}\wedge\beta^n_n)\wedge c_{n-1})\wedge((\alpha_{n+1}\wedge\beta^n_n)\wedge d_{n-1}))\vee[\cdots\vee\\ [((((\alpha_{n+1}\wedge\beta^n_2)\wedge c_1)\wedge((\alpha_{n+1}\wedge\beta^n_2)\wedge d_1))\vee\alpha_1]\cdots]]]\end{array}}\,2n\cdot\mathsf{s}}{\begin{array}{l}\bar{\alpha}_{n+1}\vee\\ [((\alpha_{n+1}\wedge c_n)\wedge(\alpha_{n+1}\wedge d_n))\vee\\ [((((\alpha_{n+1}\wedge\beta^n_n)\wedge c_{n-1})\wedge((\alpha_{n+1}\wedge\beta^n_n)\wedge d_{n-1}))\vee[\cdots\vee\\ [((((\alpha_{n+1}\wedge\beta^n_2)\wedge c_1)\wedge((\alpha_{n+1}\wedge\beta^n_2)\wedge d_1))\vee\alpha_1]\cdots]]]\end{array}}\,(2n-1)\cdot\mathsf{c}{\downarrow}\,,$$

where we use macro inference rules as explained in Remark 2.22. (Example 2.14 explains the central step in the preceding derivation.) The formulae appearing in the middle of the previous derivation are the largest. Since $|\alpha_{n+1}| = 2$ and $|\beta^n_k| = 2(n-k+1)$, their size is $2n\cdot 2+6+2(3n-3+\sum_{k=2}^{n}|\beta^n_k|)+2 = 2n^2+8n+2$. Each $\mathsf{i}{\downarrow}$ macro inference step involves one s and two $\mathsf{ai}{\downarrow}$ steps in KS, for a total of six steps, including $=$ ones; each $\mathsf{c}{\downarrow}$ macro inference step involves one m and two $\mathsf{ac}{\downarrow}$ steps in KS, for a total of six steps. So, the length of Φ is $2n\cdot 6+2n\cdot 2+(2n-1)\cdot 6 = 28n-6$, and so $|\Phi| \leq (28n-6)(2n^2+8n+2) \in \mathsf{O}(n^3)$. □

Note that in the previous proof, by working with macro inference rules, we get a better upper bound for KS than if we worked in KSg and then applied Theorem 2.22.

Theorem 3.12. *There are* KS *proofs of Statman tautologies whose size is quadratic in the size of the tautologies they prove.*

Proof. Tautology S_1 is trivially provable by an instance of the $\mathsf{i}{\downarrow}$ macro rule. By repeatedly applying the previous lemma, we obtain proofs of all Statman tautologies S_n, whose size is $\mathsf{O}(n^4)$. Since $|\mathsf{S}_n| \in \mathsf{O}(n^2)$ (see Remark 3.9), the statement follows. □

This is enough to conclude that analytic $\mathsf{Gentzen}$ does not polynomially simulate KSg and KS. Some could argue that Statman tautologies are artificial in their forcing exponential Gentzen proofs into 'wildly' branching. However, notice that both notions of proof and analyticity in Gentzen systems 'get into tautologies' from the outside inwards. In other words, the restricted notion of analyticity in Gentzen systems is strongly correlated to the restricted notion of proof that leads to exponential-size proofs. In CoS, both notions are more liberal, to the advantage of proof complexity. We pay a price for this in terms of proof-search complexity: there is research aimed at improving the situation, with very promising results; see [Kah06a].

We note that polynomial proofs on Statman tautologies are obtained by a very small dose of deep inference. In fact, the trick is done by the switch and interaction instances

in the proof in Lemma 3.11: they all operate just below the 'surface' of a formula. This leads us to state a currently open problem, in Section 6.6.

4. Calculus of Structures and Frege Systems

In this section, we prove that CoS and Frege systems are p-equivalent.

In the following definitions about Frege systems, we do not assume that the language of formulae coincides with the CoS one, but, as always, this is not a very important issue.

Definition 4.1. Given a language of propositional logic formulae built over a complete base of connectives, a *Frege* (*proof*) *system* is a finite collection of sound *inference rules*, each of which is a tuple of $n > 0$ formulae such that from $n-1$ *premisses* one *conclusion* is derived; inference rules with 0 premisses are called *axioms*. Given a Frege system, a *Frege derivation* of *length* l with *premisses* $\alpha_1, \ldots, \alpha_h$ and *conclusion* β_l is a sequence of formulae $\beta_1, \ldots, \beta_l$, such that each β_i either belongs to $\{\alpha_1, \ldots, \alpha_h\}$ or is the conclusion of an instance of an inference rule whose premisses belong to $\beta_1, \ldots, \beta_{i-1}$, where $1 \leq i \leq l$; a *Frege proof* of β is a Frege derivation with no premisses and conclusion β; we use Υ for derivations. We require of each Frege system to be *implicationally complete*, *i.e.*, whenever $(\alpha_1 \wedge \cdots \wedge \alpha_h) \rightarrow \beta$ is valid there is a derivation with premisses $\alpha_1, \ldots, \alpha_h$ and conclusion β in the proof system. The *size* of a Frege derivation Υ is the number of unit, atom, and variable occurrences that it contains, and is indicated by $|\Upsilon|$.

The following 'robustness' theorem can easily be proved.

Theorem 4.2. (*Robustness*, Cook-Reckhow, [CR79]) *All Frege systems in the same language are p-equivalent.*

The theorem has been generalised by Reckhow to Frege systems in any language (under 'natural translations') [Rec76], but we do not need this level of generality in our article. The robustness theorem allows us to work with just one Frege system, and we arbitrarily choose the following, taken from [Bus87] and modified by adding axioms F_{14}, F_{15}, F_{16}, and F_{17} in order to deal with units.

Definition 4.3. Frege system Frege, over the language of formulae freely generated by units, non-negated formula variables, and the connectives $\vee$, $\wedge$, $\rightarrow$, and $\neg$, has inference rules as shown in Figure 5, where the formulae $\mathsf{F}_1, \ldots, \mathsf{F}_{17}$ are axioms and the inference rule mp is called *modus ponens*.

Remark 4.4. In the following, every SKS formula is implicitly translated into a Frege formula in the obvious way, and vice versa; in particular, we translate Frege's formulae of the kind $\alpha \rightarrow \beta$ into SKS formulae $\bar{\alpha} \vee \beta$.

Remark 4.5. In Frege systems, distinguishing atoms from formula variables is unnecessary because we only instantiate rules by general substitution (as opposed to renaming). So, from now on, we assume that CoS atoms correspond to Frege formula variables. Since in system Frege we have a connective for negation, we can also assume that when dual atoms and formula variables appear in an SKSg formula, their Frege translation only uses $\neg$; for example, the SKSg formula $[A \vee \bar{A}] \wedge [a \vee \bar{a}]$ is translated into Frege formula $[A \vee \neg A] \wedge [B \vee \neg B]$ or $[A \vee \neg A] \wedge [\neg B \vee B]$ or $[\neg A \vee A] \wedge [B \vee \neg B]$ or $[\neg A \vee A] \wedge [\neg B \vee B]$. Conversely, Frege formula $[A \vee \neg A] \wedge [B \vee \neg B]$ is translated into SKSg or SKS formula $[a \vee \bar{a}] \wedge [b \vee \bar{b}]$ or $[a \vee \bar{a}] \wedge [\bar{b} \vee b]$ or $[\bar{a} \vee a] \wedge [b \vee \bar{b}]$ or $[\bar{a} \vee a] \wedge [\bar{b} \vee b]$ or one such formula with formula variables in the place of some of the atoms. As always, we use atoms when we need to use SKS atomic structural rules, we use formula variables when we need to instantiate formulae and derivations, and otherwise we can choose both.

Translating Frege into SKSg derivations is straightforward, given that the cut rule of SKSg can easily simulate modus ponens.

Axioms:

$\mathsf{F}_1 \equiv A \to (B \to (A \wedge B))$
$\mathsf{F}_2 \equiv (A \wedge B) \to A$
$\mathsf{F}_3 \equiv (A \wedge B) \to B$
$\mathsf{F}_4 \equiv A \to [A \vee B]$
$\mathsf{F}_5 \equiv B \to [A \vee B]$
$\mathsf{F}_6 \equiv \neg\neg A \to A$
$\mathsf{F}_7 \equiv A \to \neg\neg A$
$\mathsf{F}_8 \equiv A \to (B \to A)$
$\mathsf{F}_9 \equiv \neg A \to (A \to B)$
$\mathsf{F}_{10} \equiv (A \to (B \to C)) \to ((A \to B) \to (A \to C))$
$\mathsf{F}_{11} \equiv (A \to C) \to ((B \to C) \to ([A \vee B] \to C))$
$\mathsf{F}_{12} \equiv (A \to (B \to C)) \to (B \to (A \to C))$
$\mathsf{F}_{13} \equiv (A \to B) \to (\neg B \to \neg A)$
$\mathsf{F}_{14} \equiv \mathsf{f} \to (A \wedge \neg A)$
$\mathsf{F}_{15} \equiv (A \wedge \neg A) \to \mathsf{f}$
$\mathsf{F}_{16} \equiv \mathsf{t} \to [A \vee \neg A]$
$\mathsf{F}_{17} \equiv [A \vee \neg A] \to \mathsf{t}$

Inference rule:

$$\mathsf{mp}\,\frac{A \quad A \to B}{B}$$

FIGURE 5. System Frege.

Theorem 4.6. *For every* Frege *derivation* Υ *with premisses* $\alpha_1, \ldots, \alpha_h$, *where* $h \geq 0$, *and conclusion* β, *there is a derivation* $\begin{array}{c}\alpha_1 \wedge \cdots \wedge \alpha_h \\ \Phi \Big\| \mathsf{SKSg} \\ \beta\end{array}$ *; if* l *and* n *are, respectively, the length and size of* Υ, *then the length and size of* Φ *are, respectively,* $\mathsf{O}(l)$ *and* $\mathsf{O}(n^2)$.

Proof. The axioms F_i of Frege are tautologies, so each one has a proof Φ_i in SKSg, for $1 \leq i \leq 17$; for example F_1 and F_{10} are, respectively, proved by

$$\Phi_1 = \begin{array}{c}\mathsf{i}\!\downarrow \dfrac{\mathsf{t}}{[\bar{A} \vee \bar{B}] \vee (A \wedge B)} \\ = \dfrac{}{\bar{A} \vee [\bar{B} \vee (A \wedge B)]}\end{array} \qquad \text{and} \qquad \Phi_{10} = \begin{array}{c}\mathsf{i}\!\downarrow \dfrac{\mathsf{i}\!\downarrow \dfrac{\mathsf{t}}{(A \wedge (B \wedge \bar{C})) \vee [\bar{A} \vee [\bar{B} \vee C]]}}{(A \wedge (B \wedge \bar{C})) \vee [\bar{A} \vee [(\bar{B} \wedge \mathsf{t}) \vee C]]} \\ \mathsf{s}\,\dfrac{\mathsf{i}\!\downarrow \dfrac{}{(A \wedge (B \wedge \bar{C})) \vee [\bar{A} \vee [(\bar{B} \wedge [A \vee \bar{A}]) \vee C]]}}{(A \wedge (B \wedge \bar{C})) \vee [\bar{A} \vee [[(\bar{B} \wedge A) \vee \bar{A}] \vee C]]} \\ \mathsf{c}\!\downarrow \dfrac{}{(A \wedge (B \wedge \bar{C})) \vee [(A \wedge \bar{B}) \vee [[\bar{A} \vee \bar{A}] \vee C]]}\end{array} .$$

We proceed by induction on the length of $\Upsilon = \beta_1, \ldots, \beta_k, \beta$ and we prove the existence of a derivation $\begin{array}{c}\alpha_1 \wedge \cdots \wedge \alpha_h \\ \Phi' \Big\| \mathsf{SKSg} \\ (\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta\end{array}$. The base case $k = 0$ is as follows: 1) if β is a premiss, then $\Phi' = \beta$; 2) if $\beta \equiv \mathsf{F}_i\sigma$, for some i and σ, then $\Phi' = \Phi_i\sigma$. For the inductive step, given $\Upsilon_k = \beta_1, \ldots, \beta_k$ and $\begin{array}{c}\gamma_k \\ \Phi'_k \Big\| \mathsf{SKSg} \\ \beta_1 \wedge \cdots \wedge \beta_k\end{array}$, where γ_k is the conjunction of premisses of Υ_k, we consider the following cases:

- if β is a premiss, then $\Phi' = \begin{array}{c}\gamma_k \wedge \beta \\ \Phi'_k \wedge \beta \Big\| \mathsf{SKSg} \\ (\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta\end{array}$;

- if $\beta \equiv \mathsf{F}_i\sigma$, for some i and substitution σ, then $\Phi' = \quad = \dfrac{\begin{array}{c}\gamma_k\\ \Phi'_k \Big\| \mathsf{SKSg}\\ \beta_1 \wedge \cdots \wedge \beta_k\end{array}}{\begin{array}{c}(\beta_1 \wedge \cdots \wedge \beta_k) \wedge \mathsf{t}\\ (\beta_1 \wedge \cdots \wedge \beta_k) \wedge \Phi_i\sigma \Big\| \mathsf{SKSg}\\ (\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta\end{array}}$;
- if β is the conclusion of an instance $\mathsf{mp}\dfrac{\beta_{k'} \quad \beta_{k'} \to \beta}{\beta}$, where $\beta_{k''} \equiv \beta_{k'} \to \beta$ and $1 \leq k', k'' \leq k$, then

$$\Phi' = \quad = \dfrac{\begin{array}{c}\mathsf{c}\uparrow \dfrac{\mathsf{c}\uparrow \dfrac{\begin{array}{c}\gamma_k\\ \Phi'_k \Big\| \mathsf{SKSg}\\ \beta_1 \wedge \cdots \wedge \beta_{k''} \wedge \cdots \wedge \beta_k\end{array}}{\beta_1 \wedge \cdots \wedge \beta_{k'} \wedge \cdots \wedge (\beta_{k''} \wedge \beta_{k''}) \wedge \cdots \wedge \beta_k}}{\beta_1 \wedge \cdots \wedge (\beta_{k'} \wedge \beta_{k'}) \wedge \cdots \wedge (\beta_{k''} \wedge \beta_{k''}) \wedge \cdots \wedge \beta_k}\end{array}}{\begin{array}{c}\mathsf{s}\dfrac{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge (\beta_{k'} \wedge [\bar{\beta}_{k'} \vee \beta])}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge [(\beta_{k'} \wedge \bar{\beta}_{k'}) \vee \beta]}\\ \mathsf{i}\uparrow \dfrac{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge [(\beta_{k'} \wedge \bar{\beta}_{k'}) \vee \beta]}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge [\mathsf{f} \vee \beta]}\end{array}}{(\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta}\quad ,$$

where, without loss of generality, we assumed $k' < k''$.

At every inductive step the length of the SKSg derivation is only increased by an $\mathsf{O}(1)$ number of inference steps. From $\begin{array}{c}\alpha_1 \wedge \cdots \wedge \alpha_h\\ \Phi' \Big\| \mathsf{SKSg}\\ (\beta_1 \wedge \cdots \wedge \beta_k) \wedge \beta\end{array}$ we can obtain the desired derivation $\begin{array}{c}\alpha_1 \wedge \cdots \wedge \alpha_h\\ \Phi \Big\| \mathsf{SKSg}\\ \beta\end{array}$ by applying once the $\mathsf{w}\uparrow$ rule. So, the length of Φ is $\mathsf{O}(k)$. From this, and after inspecting the aforesaid derivations, it follows that $|\Phi| \in \mathsf{O}(k^2 m)$, where m is the maximum size of a formula appearing in Υ, and so $|\Phi| \in \mathsf{O}(n^2)$, where $|\Upsilon| = n$. □

Corollary 4.7. SKSg *and* SKS *p-simulate* Frege.

Proof. The statement for SKSg follows from Theorem 4.6, and that for SKS from this and Corollary 2.23. □

Translating derivations from SKSg to Frege requires more effort than the converse, partly because of the need to simulate deep inference, and partly because of the large 'amount of inference' of =-rule instances. The next two lemmas take care of these two issues.

Lemma 4.8. *For every* SKS *context* $\xi\{\ \}$ *and formulae* α *and* β*, there is a* Frege *derivation with premiss* $\alpha \to \beta$ *and conclusion* $\xi\{\alpha\} \to \xi\{\beta\}$ *whose length is* $\mathsf{O}(m)$ *and size is* $\mathsf{O}(n^2)$*, where* $m = |\xi\{\ \}|$ *and* $n = |\xi\{\alpha\} \to \xi\{\beta\}|$.

Proof. Consider four Frege proofs Υ', Υ'', Υ''', and Υ'''', respectively of the four tautologies

$$(A \to B) \to ([A \vee C] \to [B \vee C]) \quad , \qquad (A \to B) \to ([C \vee A] \to [C \vee B]) \quad ,$$
$$(A \to B) \to ((A \wedge C) \to (B \wedge C)) \quad , \qquad (A \to B) \to ((C \wedge A) \to (C \wedge B)) \quad .$$

We proceed by induction on the structure of $\xi\{\ \}$. If $\xi\{\ \} \equiv \xi_1\{\{\ \} \vee \gamma_1\}$, we build Frege derivation

$$\Upsilon_1 = \alpha \to \beta, \Upsilon'\{A/\alpha, B/\beta, C/\gamma_1\}, \alpha \vee \gamma_1 \to \beta \vee \gamma_1 \quad ;$$

we build Υ_1 similarly if $\xi\{\ \} \equiv \xi_1\{\{\ \} \wedge \gamma_1\}$ or $\xi\{\ \} \equiv \xi_1\{\gamma_1 \vee \{\ \}\}$ or $\xi\{\ \} \equiv \xi_1\{\gamma_1 \wedge \{\ \}\}$. Given $\xi_1\{\ \} \equiv \xi_2\{\{\ \} \vee \gamma_2\}$ or $\xi_1\{\ \} \equiv \xi_2\{\{\ \} \wedge \gamma_2\}$ or $\xi_1\{\ \} \equiv \xi_2\{\gamma_2 \vee \{\ \}\}$ or $\xi_1\{\ \} \equiv \xi_2\{\gamma_2 \wedge \{\ \}\}$ we build Υ_2 analogously to Υ_1, and the premiss of Υ_2 is the conclusion of Υ_1. We proceed this way until we build Υ_l, whose conclusion is $\xi\{\alpha\} \to \xi\{\beta\}$, where $l \leq m$. We obtain the desired derivation Υ by concatenating $\Upsilon_1, \ldots, \Upsilon_l$. Since the length and size of Υ', Υ'', Υ''', and Υ'''' are independent of $\xi\{\ \}$, α, and β, the length of Υ is $\mathsf{O}(m)$ and its size is $\mathsf{O}(mn)$, and so $\mathsf{O}(n^2)$. □

Lemma 4.9. *For every* SKS *formulae α and β such that $\alpha = \beta$ there is a* Frege *derivation with premiss α, conclusion β, length $\mathsf{O}(n^3)$, and size $\mathsf{O}(n^4)$, where $n = |\alpha| + |\beta|$.*

Proof. Consider the following tautologies, derived from the equations in Figure 1:

$$
\begin{array}{cc}
[A \vee B] \leftrightarrow [B \vee A] \quad , & [A \vee \mathsf{f}] \leftrightarrow A \quad , \\
(A \wedge B) \leftrightarrow (B \wedge A) \quad , & (A \wedge \mathsf{t}) \leftrightarrow A \quad , \\
[[A \vee B] \vee C] \leftrightarrow [A \vee [B \vee C]] \quad , & [\mathsf{t} \vee \mathsf{t}] \leftrightarrow \mathsf{t} \quad , \\
((A \wedge B) \wedge C) \leftrightarrow (A \wedge (B \wedge C)) \quad , & (\mathsf{f} \wedge \mathsf{f}) \leftrightarrow \mathsf{f} \quad ,
\end{array}
\tag{1}
$$

where each expression corresponds to the two tautologies obtained by orientating each double implication. Every such tautology can be proved in Frege with a constant-size proof, so every instance $\gamma \to \gamma'$ of any of these tautologies has a Frege proof of length $\mathsf{O}(1)$ and size $\mathsf{O}(m')$, where $m' = |\gamma| + |\gamma'|$. By Lemma 4.8, for every $\xi\{\ \}$ there is a derivation with premiss $\gamma \to \gamma'$ and conclusion $\xi\{\gamma\} \to \xi\{\gamma'\}$ whose length is $\mathsf{O}(m)$ and size is $\mathsf{O}(m^2)$, where $m = |\xi\{\gamma\} \to \xi\{\gamma'\}|$. By concatenating the proof and derivation so obtained, we can build a proof of $\xi\{\gamma\} \to \xi\{\gamma'\}$ whose length is $\mathsf{O}(m)$ and size is $\mathsf{O}(m^2)$. By Remark 2.9, we can build a chain of implications

$$\alpha \equiv \alpha_1 \to \cdots \to \alpha_h \equiv \delta \equiv \beta_k \to \cdots \to \beta_1 \equiv \beta \quad ,$$

where δ is a canonical form for α and β, $h + k$ is $\mathsf{O}(n^2)$, and each implication $\alpha_i \to \alpha_{i+1}$ and $\beta_{i+1} \to \beta_i$ is a tautology of the form $\xi\{\gamma\} \to \xi\{\gamma'\}$, such that $\gamma \to \gamma'$ is an instance of one of the tautologies 1. By concatenating the proofs of every $\xi\{\gamma\} \to \xi\{\gamma'\}$ by mp, we obtain a derivation with premiss α, conclusion β, length $\mathsf{O}(n^3)$, and size $\mathsf{O}(n^4)$. □

Lemma 4.10. *For every inference step $\nu\dfrac{\alpha}{\beta}$, where ν is a rule of* SKSg, *there is a* Frege *derivation with premiss α, conclusion β, length $\mathsf{O}(n)$, and size $\mathsf{O}(n^2)$, where $n = |\alpha| + |\beta|$.*

Proof. Each of the following tautologies, corresponding to the inference rules in Figure 2, can be proved in Frege with a constant-size proof:

$$
\begin{array}{ccc}
(A \wedge \neg A) \to \mathsf{f} \quad , & A \to \mathsf{t} \quad , & A \to (A \wedge A) \quad , \\
\mathsf{f} \to [A \vee \neg A] \quad , & \mathsf{f} \to A \quad , & [A \vee A] \to A \quad , \\
 & & (A \wedge [B \vee C]) \to [(A \wedge B) \vee C] \quad .
\end{array}
\tag{2}
$$

Let $\nu\dfrac{\alpha}{\beta} = \nu\dfrac{\xi\{\gamma\}}{\xi\{\delta\}}$, where $\nu\dfrac{\gamma}{\delta}$ is an instance of ν. There is a Frege proof Υ of $\gamma \to \delta$, whose length is $\mathsf{O}(1)$ and size is $\mathsf{O}(n)$, obtained by instantiating the corresponding proof to ν among those in 2. By Lemma 4.8, there exists a Frege derivation Υ' with premiss $\gamma \to \delta$, conclusion $\xi\{\gamma\} \to \xi\{\delta\}$, length $\mathsf{O}(n)$, and size $\mathsf{O}(n^2)$. By concatenating Υ and Υ' we obtain a proof Υ'' of $\xi\{\gamma\} \to \xi\{\delta\}$. From Υ'', by using mp, we obtain the desired derivation with premiss $\alpha \equiv \xi\{\gamma\}$ and conclusion $\beta \equiv \xi\{\delta\}$. □

Theorem 4.11. *For every derivation* $\begin{array}{c}\alpha\\ \Phi\|\mathsf{SKSg}\\ \beta\end{array}$ *there is a* Frege *derivation* Υ *with premiss* α *and conclusion* β*; if* n *is the size of* Φ*, then the length and size of* Υ *are, respectively,* $\mathsf{O}(n^4)$ *and* $\mathsf{O}(n^5)$.

Proof. The statement immediately follows from Lemmas 4.9 and 4.10, after assuming that the length of Φ is $\mathsf{O}(n)$. □

Corollary 4.12. Frege *p-simulates* SKSg *and* SKS.

Remark 4.13. As evidenced by the proofs of Theorems 4.6 and 4.11, it does not really matter, for establishing the p-simulations, precisely which inference rules are adopted by the CoS and Frege systems. In fact, the simulations work because the simulating systems are implicationally complete and their set of proofs is closed under substitution. This way, the constant-size proofs in one system, simulating the rules of the other system, can be instantiated at a linear cost in order to simulate instances of rules. We can then use a robustness theorem (see Theorem 4.2) for Frege in order to establish a robustness theorem for CoS, possibly also for systems on mutually different languages: given two implicationally complete CoS systems, we p-simulate each in two appropriate Frege systems and use Frege robustness.

5. Extension and Substitution

In this section, we show how CoS systems can be extended with the Tseitin extension rule and with the substitution rule, analogously to Frege systems. We also show the p-equivalence of all these systems, as described in the box of the diagram in the Introduction. As always, we operate under robustness theorems (relying on the mentioned one, Theorem 4.2) that ensure that the proof complexity properties we establish for the specific systems actually hold for the formalisms they belong to.

Definition 5.1. An *extended Frege* (*proof*) *system* is a Frege system augmented with the (*Tseitin*) *extension rule*, which is a rule with no premisses and whose instances $A \leftrightarrow \beta$ are such that the variable A does not appear before in the derivation, nor appears in β or in the conclusion of the proof. We write $A \notin \alpha$ to state that variables A and $\bar{A}$ do not appear in formula α. The symbol $\leftrightarrow$ stands for logical equivalence, and the specific syntax of the expressions $A \leftrightarrow \beta$ depends on the language of the Frege system in use. In the following, we consider $A \leftrightarrow \beta$ a shortcut for $(A \to \beta) \wedge (\beta \to A)$. We denote by xFrege the proof system where a proof is a derivation with no premisses, conclusion α_k, and shape

$$\alpha_1, \ldots, \alpha_{i_1-1}, \overbrace{A_1 \leftrightarrow \beta_1}^{\alpha_{i_1} \equiv}, \alpha_{i_1+1}, \ldots, \alpha_{i_h-1}, \overbrace{A_h \leftrightarrow \beta_h}^{\alpha_{i_h} \equiv}, \alpha_{i_h+1}, \ldots, \alpha_k \quad ,$$

where all the conclusions of extension instances $\alpha_{i_1}, \ldots, \alpha_{i_h}$ are singled out and

$$A_1 \notin \alpha_1, \ldots, \alpha_{i_1-1}, \beta_1, \alpha_k \quad , \qquad \ldots \quad , \qquad A_h \notin \alpha_1, \ldots, \alpha_{i_h-1}, \beta_h, \alpha_k \quad ,$$

and the rest of the proof is as in Frege.

Remark 5.2. We could have equivalently defined an xFrege proof of α as a Frege derivation with conclusion α and premisses $\{A_1 \leftrightarrow \beta_1, \ldots, A_h \leftrightarrow \beta_h\}$ such that $A_1, \bar{A}_1, \ldots, A_h, \bar{A}_h$ are mutually distinct and $A_1 \notin \beta_1, \alpha$ and … and $A_h \notin \beta_1, \ldots, \beta_h, \alpha$. Notice that xFrege is indeed a proof system in the sense that it proves tautologies. In fact, given the xFrege proof just mentioned, we obtain a Frege proof by applying to it, in order, the substitutions $\sigma_h = A_h/\beta_h, \ldots, \sigma_1 = A_1/\beta_1$, and by prepending to it proofs of the tautologies $\beta_1 \leftrightarrow \beta_1, (\beta_2 \leftrightarrow \beta_2)\sigma_1, \ldots, (\beta_h \leftrightarrow \beta_h)\sigma_{h-1} \cdots \sigma_1$. In general, a proof so obtained is exponentially bigger than the xFrege one it derives from.

SKSg can analogously be extended, but there is no need to create a special rule; we only need to broaden the criterion by which we recognize a proof.

Definition 5.3. An *extended* SKSg *proof* of α is an SKSg derivation with conclusion α and premiss $[\bar{A}_1 \vee \beta_1] \wedge [\bar{\beta}_1 \vee A_1] \wedge \cdots \wedge [\bar{A}_h \vee \beta_h] \wedge [\bar{\beta}_h \vee A_h]$, where $A_1, \bar{A}_1, \ldots, A_h, \bar{A}_h$ are mutually distinct and $A_1 \notin \beta_1, \alpha$ and … and $A_h \notin \beta_1, \ldots, \beta_h, \alpha$. We denote by xSKSg the proof system whose proofs are extended SKSg proofs.

Theorem 5.4. *For every* xFrege *proof of length l and size n there exists an* xSKSg *proof of the same formula and whose length and size are, respectively,* $\mathsf{O}(l)$ *and* $\mathsf{O}(n^2)$.

Proof. Consider an xFrege proof as in Definition 5.1. By Remark 5.2 and Theorem 4.6, there exists the following xSKSg proof, whose length and size are yielded by 4.6:

$$\begin{array}{c} [\bar{A}_1 \vee \beta_1] \wedge [\bar{\beta}_1 \vee A_1] \wedge \cdots \wedge [\bar{A}_h \vee \beta_h] \wedge [\bar{\beta}_h \vee A_h] \\ \Big\| \mathsf{SKSg} \\ \alpha_k \end{array} .$$

□

Although not strictly necessary to establish the equivalence of the four extended formalisms (see diagram in the Introduction), the following theorem is very easy to prove.

Theorem 5.5. *For every* xSKSg *proof of size n there exists an* xFrege *proof of the same formula and whose length and size are, respectively,* $\mathsf{O}(n^4)$ *and* $\mathsf{O}(n^5)$.

Proof. Consider an xSKSg proof as in Definition 5.3. The statement is an immediate consequence of Theorem 4.11, after observing that there is an $\mathsf{O}(h)$-length and $\mathsf{O}(hn)$-size xFrege proof

$$A_1 \leftrightarrow \beta_1, \ldots, A_h \leftrightarrow \beta_h, \ldots, (A_1 \leftrightarrow \beta_1) \wedge \cdots \wedge (A_h \leftrightarrow \beta_h) \quad .$$

□

Corollary 5.6. *Systems* xFrege *and* xSKSg *are p-equivalent.*

We now move to the substitution rule.

Definition 5.7. A *substitution Frege* (*proof*) *system* is a Frege system augmented with the *substitution rule* $\mathsf{sub}\dfrac{A}{A\sigma}$. We denote by sFrege the proof system where a proof is a derivation with no premisses, conclusion α_k, and shape

$$\alpha_1, \ldots, \alpha_{i_1-1}, \overbrace{\alpha_{j_1}\sigma_1}^{\alpha_{i_1} \equiv}, \alpha_{i_1+1}, \ldots, \alpha_{i_h-1}, \overbrace{\alpha_{j_h}\sigma_h}^{\alpha_{i_h} \equiv}, \alpha_{i_h+1}, \ldots, \alpha_k \quad ,$$

where all the conclusions of substitution instances $\alpha_{i_1}, \ldots, \alpha_{i_h}$ are singled out, $\alpha_{j_1} \in \{\alpha_1, \ldots, \alpha_{i_1-1}\}, \ldots, \alpha_{j_h} \in \{\alpha_1, \ldots, \alpha_{i_h-1}\}$, and the rest of the proof is as in Frege.

We rely on the following result.

Theorem 5.8. (Cook-Reckhow and Krajíček-Pudlák, [CR79, KP89]) *Systems* xFrege *and* sFrege *are p-equivalent.*

We can extend SKSg with the same substitution rule as for Frege. The rule is used like other proper rules of system SKSg, so its instances are interleaved with $=$-rule instances.

Definition 5.9. An sSKSg proof is a proof of SKSg where, in addition to the inference steps generated by rules of SKSg, we admit inference steps obtained as instances of the *substitution rule* $\mathsf{sub}\dfrac{A}{A\sigma}$.

This rule does not fit any of the usual deep-inference rule classes (see Section 2), and (as in Frege systems) is not sound, in the sense that the premiss does not imply the conclusion. However, of course, if the premiss is provable the conclusion also is.

Remark 5.10. Notice that instances of the substitution rule cannot be used inside a context; for example, the expression on the left is not a valid sSKSg proof, while the one on the right is:

$$\mathsf{sub?}\,\frac{\mathsf{i}{\downarrow}\,\dfrac{\mathsf{t}}{A\vee\bar{A}}}{(B\wedge C)\vee\bar{A}}\quad,\qquad \mathsf{sub}\,\frac{\mathsf{i}{\downarrow}\,\dfrac{\mathsf{t}}{A\vee\bar{A}}}{(B\wedge C)\vee[\bar{B}\vee\bar{C}]}\quad.$$

In the so-called 'Formalism B' of deep inference, which is currently under development [Gug04], and for which all the proof-complexity results in this article apply unchanged, substitution becomes part of the composition mechanism of proofs, rather than an odd extension to the set of rules.

For the time being, we can establish the promised p-equivalence of all extended systems by completing the diagram in the Introduction with the last two missing steps.

Theorem 5.11. *For every* xSKSg *proof of size n there exists an* sSKSg *proof of the same formula and whose length and size are, respectively,* $\mathsf{O}(n)$ *and* $\mathsf{O}(n^2)$.

Proof. Consider the xSKSg proof

$$\begin{array}{c}[\bar{A}_1\vee\beta_1]\wedge[\bar{\beta}_1\vee A_1]\wedge\cdots\wedge[\bar{A}_h\vee\beta_h]\wedge[\bar{\beta}_h\vee A_h]\\ \Phi\,\Big\|\,\mathsf{SKSg}\\ \alpha\end{array}\quad,\qquad\text{where}\qquad\begin{array}{l}A_1\notin\beta_1,\alpha\quad,\\ \dots\quad,\\ A_h\notin\beta_1,\dots,\beta_h,\alpha\quad,\end{array}$$

and let us call its premiss γ. We can build the sSKSg proof

$$\begin{array}{c}\mathsf{i}{\downarrow}\,\dfrac{\mathsf{t}}{\bar{\gamma}\vee\gamma}\\ \bar{\gamma}\vee\Phi\,\Big\|\,\mathsf{SKSg}\\ \mathsf{sub}\,\dfrac{[(A_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{A}_1)\vee\cdots\vee(A_h\wedge\bar{\beta}_h)\vee(\beta_h\wedge\bar{A}_h)]\vee\alpha}{\mathsf{c}{\downarrow}\,\dfrac{[(A_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{A}_1)\vee\cdots\vee(\beta_h\wedge\bar{\beta}_h)\vee(\beta_h\wedge\bar{\beta}_h)]\vee\alpha}{\mathsf{i}{\uparrow}\,\dfrac{[(A_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{A}_1)\vee\cdots\vee(\beta_h\wedge\bar{\beta}_h)]\vee\alpha}{=\dfrac{[(A_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{A}_1)\vee\cdots\vee\mathsf{f}]\vee\alpha}{\dfrac{\vdots}{=\dfrac{\mathsf{sub}\,\dfrac{[(A_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{A}_1)]\vee\alpha}{\mathsf{c}{\downarrow}\,\dfrac{[(\beta_1\wedge\bar{\beta}_1)\vee(\beta_1\wedge\bar{\beta}_1)]\vee\alpha}{\mathsf{i}{\uparrow}\,\dfrac{(\beta_1\wedge\bar{\beta}_1)\vee\alpha}{=\dfrac{\mathsf{f}\vee\alpha}{\alpha}}}}}}}}}}\quad.$$

□

Corollary 5.12. sSKSg *p-simulates* xSKSg.

Theorem 5.13. *For every* sSKSg *proof of size n there exists a proof of the same formula in* sFrege, *whose length and size are, respectively,* $\mathsf{O}(n^4)$ *and* $\mathsf{O}(n^5)$.

Proof. Every sSKSg proof has shape

$$\begin{array}{c}
\mathsf{t} \\
\Phi_0 \,\Big\|\, \mathsf{SKSg} \\
\mathsf{sub}\,\dfrac{\alpha_1}{\alpha_1\sigma_1} \\
\Phi_1 \,\Big\|\, \mathsf{SKSg} \\
\vdots \\
\Phi_{h-1} \,\Big\|\, \mathsf{SKSg} \\
\mathsf{sub}\,\dfrac{\alpha_h}{\alpha_h\sigma_h} \\
\Phi_h \,\Big\|\, \mathsf{SKSg} \\
\alpha_{h+1}
\end{array} .$$

By Theorem 4.11, for each of $\Phi_0, \ldots, \Phi_h$ there exist Frege derivations $\Upsilon_0, \ldots, \Upsilon_h$ with the same premiss and conclusion, respectively. We can then build the proof

$$\overbrace{\ldots,\alpha_1}^{\Upsilon_0},\overbrace{\alpha_1\sigma_1,\ldots}^{\Upsilon_1},\ldots,\overbrace{\ldots,\alpha_h}^{\Upsilon_{h-1}},\overbrace{\alpha_h\sigma_h,\ldots,\alpha_{h+1}}^{\Upsilon_h}$$

in sFrege; the cited theorem also yields its length and size. □

Corollary 5.14. sFrege *p-simulates* sSKSg.

Nothing prevents us from using Tseitin extension and the substitution rule with system SKS, or any other atomic or nonatomic CoS system. The integration of these mechanisms into CoS is similar to their integration into Frege systems, as the simplicity of the arguments showing p-equivalence testifies.

6. Open Problems

We conclude the article with a list of open problems, some of which are currently investigated by us and other researchers.

6.1. **Relation with Resolution and Other Formalisms.** In this article, we explored the relation between CoS and Frege systems, and in the cited literature the relation between CoS and Gentzen systems has been explored in depth. There are, of course, other formalisms, like resolution, whose relation with CoS might lead to some interesting research directions. For example, the note [Gug03] shows how simply, compared to Gentzen systems, KS expresses resolution (analytically, of course).

6.2. **Does Cocontraction Provide for an Exponential Speedup?** As we argued in Remark 3.6, we do not know whether KSg p-simulates KSg$\cup\{\mathsf{c}\uparrow\}$, or, equivalently, whether KS p-simulates KS$\cup\{\mathsf{ac}\uparrow\}$.

Our intuition, as well as some clues, like the mutual behaviour of the 'atomic flows' of contraction and cocontraction (see [GG08]) would lead us to believe that cocontraction indeed provides for an exponential speedup. However, we know that in similar situations, like for dag-like versus tree-like Frege systems, intuition was fallacious.

If cocontraction yields an exponential speedup, we obtain an even stronger analytic system than KSg, which is, in turn, stronger than analytic Gentzen. This would draw interest to a hierarchy of analytic proof systems of different strength.

Unless we prove the p-equivalence of KSg and KSg$\cup\{\mathsf{c}\uparrow\}$, we tend to consider cocontraction a simple rule-based mechanism for compressing proofs, like cut, extension, and substitution.

6.3. **Pigeonhole in Analytic CoS.** Does the pigeonhole principle, in particular in its relational variety, admit polynomially growing proofs in KS? If not, does it in KS$\cup\{\mathsf{ac}\uparrow\}$?

Investigating this problem could be relevant to the more general, following one (Problem 6.4), about the ability of analytic CoS to simulate CoS, and so Frege. In fact, the pigeonhole principle generates some of the hardest classes of tautologies known.

We note that in [Jap08], Japaridze shows polynomially growing proofs for the pigeonhole class of tautologies in a deep-inference system over certain circuit-like sequents, called 'cirquents'. In this case, the speedup is obtained by the sharing of logical expressions in circuits.

In [Jeř09], Jeřábek shows that there are polynomial-time constructible proofs in KS$\cup\{\mathsf{ac}\uparrow\}$ of the functional and onto variants of the pigeonhole principle.

6.4. **Relative Strength of Analytic CoS and CoS.** Some recent major progress has been made in [Jeř09]. There, Jeřábek uses a construction on threshold formulae in the monotone sequent calculus, by Atserias, Galesi and Pudlák [AGP02], to show that analytic CoS quasipolynomially simulates CoS. In [BGGP09], we provide a direct and simplified construction based on atomic flows [GG08] and threshold formulae.

Because of these recent advances, we expect that analytic CoS p-simulates CoS. A more in-depth discussion of this subject is in [BGGP09]. If analytic CoS p-simulates CoS, then there are polynomially growing proofs of the pigeonhole principle in analytic CoS, though not necessarily in KS.

We think that investigating this problem will help us to better understand analyticity, in order to obtain for it a general and more useful definition than the one we have now. We feel that the current notion is not satisfactory because it depends on the formalism and must be defined by resorting to the syntactic structure of inference rules (or, worse, by indicating which rules are analytic and which are not).

For example, a more general, nonsyntactic definition of analyticity could be the following: a rule is analytic if, given an instance of its conclusion, the set of possible instances of the premiss is finite (this is what we call a finitely generating rule in Section 2). In this sense, an atomic 'finitary' cut rule $\mathsf{fai}\uparrow\dfrac{\xi\{a\wedge\bar{a}\}}{\xi\{\mathsf{t}\}}$, such that a appears in $\xi\{\ \}$, would be analytic. However, [BG04] shows that we can easily transform proofs in SKS into smaller- or equal-size proofs that only use $\mathsf{fai}\uparrow$ wherever $\mathsf{ai}\uparrow$ was used. So, we could deem $\mathsf{fai}\uparrow$ an analytic rule, and the system obtained from SKS by substituting $\mathsf{ai}\uparrow$ with $\mathsf{fai}\uparrow$ an analytic one, and we could immediately conclude that analytic CoS p-simulates CoS. This 'solution', however, is way too cheap.

We prefer to think that $\mathsf{fai}\uparrow$ is not an analytic rule, in some sense to be made precise. A possible point of attack is offered by the fact that $\mathsf{fai}\uparrow$ is not a local rule: it requires checking that a appears in its context, whose size is unbounded (see Remark 2.17). So, we think it could be productive to look for a notion of analyticity that is based on boundedness instead of finiteness, and tackle the separation problem between analytic CoS and CoS under that notion. The note [BG07] further explores this direction, but much more work is necessary.

6.5. **Strength of Analytic CoS Systems Plus Substitution.** We showed that CoS and Frege systems are p-equivalent, and both remain p-equivalent when extended either with Tseitin extension or substitution. However, CoS is more flexible than Frege, because it allows to 'switch off' two mechanisms that potentially provide for an exponential compression of proofs: cut and cocontraction (see Problem 6.2).

It might be interesting to study the relative strength of systems obtained by removing from SKS$\cup\{\mathsf{sub}\}$ either $\mathsf{ai}\uparrow$ or $\mathsf{ac}\uparrow$ or both. (Rule $\mathsf{aw}\uparrow$ can also be removed, but we do not see a crucial role for it.) Notice that systems KS$\cup\{\mathsf{sub}\}$ and KS$\cup\{\mathsf{ac}\uparrow,\mathsf{sub}\}$ could be considered, in some sense, analytic, and we do not know their relative strength.

6.6. **Speedup of Deep Inference Over Any Bounded-Depth System.** We saw in Theorem 3.12 that analytic CoS exhibits an exponential speedup over analytic Gentzen, for Statman tautologies. We argued that, in this case, the speedup is obtained by a rather trivial use of deep inference, because the depth at which inference has to be performed, in order to get the speedup, is constant. So, a natural question is whether there exists a class of tautologies that requires full-fledged deep inference in order to obtain efficient proofs. We think we found such a class, which is defined as follows.

Consider, for every propositional formula α, the following set of second-order formulae, for $n > 0$:

$$g(1,\alpha) \equiv \forall\beta.[((\alpha\wedge\beta)\wedge\beta)\vee(\bar{\beta}\wedge\bar{\beta})] \quad ,$$
$$g(n+1,\alpha) \equiv \forall\beta.[(g(n,\alpha\wedge\beta)\wedge g(n,\beta))\vee(g(n,\bar{\beta})\wedge g(n,\bar{\beta}))] \quad .$$

By using these formulae as a template, we can generate a set of first-order formulae, where the (complex) management of indices ensures their uniqueness:

Definition 6.1. Consider, for $m,n \geq 0$

$$h(1,m,\alpha) \equiv ((\alpha\wedge\beta_{m+1})\wedge\beta_{m+1})\vee(\bar{\beta}_{m+1}\wedge\bar{\beta}_{m+1}) \quad ,$$
$$h(n+2,m,\alpha) \equiv (h(n+1,m,\alpha\wedge\beta_{5^{n+1}+m})\wedge h(n+1,5^n+m,\beta_{5^{n+1}+m}))\vee$$
$$(h(n+1,2\cdot 5^n+m,\bar{\beta}_{5^{n+1}+m})\wedge h(n+1,3\cdot 5^n+m,\bar{\beta}_{5^{n+1}+m})) \quad .$$

Consider now

$$f(n) \equiv h(n,0,\mathsf{t}) \quad , \qquad \text{for } n > 0 \quad .$$

We define the set $\mathsf{DT} = \{f(n) \mid n > 0\}$.

The program [Gug07a] can help in understanding the nature of these formulae.

Remark 6.2. It is not difficult to verify that DT contains tautologies possessing analytic CoS proofs that grow polynomially in the size of the tautologies.

The analytic CoS proofs of the DT tautologies, when read bottom-up, work by applying interactions starting from the deepest subformulae. When this cannot be the case, we conjecture that the size of the proofs grows exponentially.

Definition 6.3. The *and/or depth* of a formula is the maximum number of alternations of conjunctions and disjunctions in the formula tree; the *and/or depth* of a context $\xi\{\ \}$ is the number of alternations of conjunctions and disjunctions between the hole and the root of the context tree. We define a *bounded-depth* CoS proof system as a CoS proof system whose inference rules only generate inference steps at a bounded depth, namely inference steps $v\dfrac{\xi\{\gamma\}}{\xi\{\delta\}}$ are such that, if $v\dfrac{\gamma}{\delta}$ is a rule instance then the and/or depth of $\xi\{\ \}$ is bounded by a given constant, and the same restriction holds for the contexts in the context closure condition of relation $=$.

Remark 6.4. Note that the nonatomic rules interaction (identity), cointeraction (cut), contraction and cocontraction require establishing duality or identity of formulae of unbounded and/or depth. So, their adoption might be considered an implicit use of deep inference. However, the atomic counterparts of these rules do not suffer this problem because the 'deep checking' is delegated to the inference mechanism. For this reason, proving the following conjecture is better done in the analytic part of system SKS.

Conjecture 6.5. In any analytic bounded-depth CoS proof system, the tautologies in DT only have proofs that grow exponentially in their size.

7. Conclusion

In this article, we showed that the calculus of structures (CoS) has the same characteristics of the Frege formalism in terms of proof complexity, including when extended with Tseitin extension and substitution.

We know that, contrary to Frege, CoS has a rich proof theory, and its proof systems enjoy several properties, arguably relevant to proof complexity, that cannot be observed in other formalisms, like locality for all inference rules. We also know that other logics, like modal logics, enjoy simple and modular presentations in deep inference, which should help in proof complexity investigations. This article establishes the basic connection between proof theory in deep inference and proof complexity.

As a consequence of its flexibility in inference rule design, CoS admits a notion of analyticity that is more flexible than its counterpart for Gentzen systems. We can then explore the strength of analytic systems in finer detail than possible in Gentzen systems. In this article, we moved forward the boundary between polynomial and exponential analytic proofs by proving Statman tautologies with polynomial, analytic deep-inference proofs.

We included a list of open problems and currently active research directions.

Acknowledgements. We thank Tom Gundersen and Ozan Kahramanoğulları for having carefully read the manuscript and for having suggested several improvements.

References

[AGP02] Albert Atserias, Nicola Galesi, and Pavel Pudlák. Monotone simulations of non-monotone proofs. *Journal of Computer and System Sciences*, 65(4):626–638, 2002.

[BG04] Kai Brünnler and Alessio Guglielmi. A first order system with finite choice of premises. In Vincent Hendricks, Fabian Neuhaus, Stig Andur Pedersen, Uwe Scheffler, and Heinrich Wansing, editors, *First-Order Logic Revisited*, Logische Philosophie, pages 59–74. Logos Verlag, 2004. `http://www.iam.unibe.ch/~kai/Papers/FinitaryFOL.pdf`.

[BG07] Paola Bruscoli and Alessio Guglielmi. On analytic inference rules in the calculus of structures. `http://cs.bath.ac.uk/ag/p/Onan.pdf`, 2007.

[BGGP09] Paola Bruscoli, Alessio Guglielmi, Tom Gundersen, and Michel Parigot. Quasipolynomial normalisation in deep inference via atomic flows and threshold formulae. Submitted. `http://cs.bath.ac.uk/ag/p/QuasiPolNormDI.pdf`, 2009.

[BL05] Kai Brünnler and Stéphane Lengrand. On two forms of bureaucracy in derivations. In Paola Bruscoli, François Lamarche, and Charles Stewart, editors, *Structures and Deduction*, pages 69–80. Technische Universität Dresden, 2005. ICALP Workshop. ISSN 1430-211X. `http://www.iam.unibe.ch/~kai/Papers/sd05.pdf`.

[Bru02] Paola Bruscoli. A purely logical account of sequentiality in proof search. In Peter J. Stuckey, editor, *Logic Programming, 18th International Conference*, volume 2401 of *Lecture Notes in Computer Science*, pages 302–316. Springer-Verlag, 2002. `http://cs.bath.ac.uk/pb/bvl/bvl.pdf`.

[Brü03a] Kai Brünnler. Atomic cut elimination for classical logic. In M. Baaz and J. A. Makowsky, editors, *CSL 2003*, volume 2803 of *Lecture Notes in Computer Science*, pages 86–97. Springer-Verlag, 2003. `http://www.iam.unibe.ch/~kai/Papers/ace.pdf`.

[Brü03b] Kai Brünnler. Two restrictions on contraction. *Logic Journal of the IGPL*, 11(5):525–529, 2003. `http://www.iam.unibe.ch/~kai/Papers/RestContr.pdf`.

[Brü04] Kai Brünnler. *Deep Inference and Symmetry in Classical Proofs*. Logos Verlag, Berlin, 2004. `http://www.iam.unibe.ch/~kai/Papers/phd.pdf`.

[Brü06a] Kai Brünnler. Cut elimination inside a deep inference system for classical predicate logic. *Studia Logica*, 82(1):51–71, 2006. `http://www.iam.unibe.ch/~kai/Papers/q.pdf`.

[Brü06b] Kai Brünnler. Deep inference and its normal form of derivations. In Arnold Beckmann, Ulrich Berger, Benedikt Löwe, and John V. Tucker, editors, *Computability in Europe 2006*, volume 3988 of *Lecture Notes in Computer Science*, pages 65–74. Springer-Verlag, July 2006. `http://www.iam.unibe.ch/~kai/Papers/n.pdf`.

[Brü06c] Kai Brünnler. Deep sequent systems for modal logic. In Guido Governatori, Ian Hodkinson, and Yde Venema, editors, *Advances in Modal Logic*, volume 6, pages 107–119. College Publications,

2006.
http://www.aiml.net/volumes/volume6/Bruennler.ps.

[Brü06d] Kai Brünnler. Locality for classical logic. *Notre Dame Journal of Formal Logic*, 47(4):557–580, 2006.
http://www.iam.unibe.ch/~kai/Papers/LocalityClassical.pdf.

[BT01] Kai Brünnler and Alwen Fernanto Tiu. A local system for classical logic. In R. Nieuwenhuis and A. Voronkov, editors, *LPAR 2001*, volume 2250 of *Lecture Notes in Artificial Intelligence*, pages 347–361. Springer-Verlag, 2001.
http://www.iam.unibe.ch/~kai/Papers/lcl-lpar.pdf.

[Bus87] Samuel R. Buss. Polynomial size proofs of the propositional pigeonhole principle. *Journal of Symbolic Logic*, 52(4):916–927, 1987.

[CK02] Peter Clote and Evangelos Kranakis. *Boolean Functions and Computation Models*. Springer-Verlag, 2002.

[CR74] Stephen Cook and Robert Reckhow. On the lengths of proofs in the propositional calculus (preliminary version). In *Proceedings of the 6th annual ACM Symposium on Theory of Computing*, pages 135–148. ACM Press, 1974.

[CR79] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, 1979.

[DG04] Pietro Di Gianantonio. Structures for multiplicative cyclic linear logic: Deepness vs cyclicity. In J. Marcinkowski and A. Tarlecki, editors, *CSL 2004*, volume 3210 of *Lecture Notes in Computer Science*, pages 130–144. Springer-Verlag, 2004.
http://www.dimi.uniud.it/~pietro/papers/Soft-copy-ps/scll.ps.gz.

[GG08] Alessio Guglielmi and Tom Gundersen. Normalisation control in deep inference via atomic flows. *Logical Methods in Computer Science*, 4(1:9):1–36, 2008.
http://arxiv.org/pdf/0709.1205.

[Gir87] Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[GS01] Alessio Guglielmi and Lutz Straßburger. Non-commutativity and MELL in the calculus of structures. In L. Fribourg, editor, *CSL 2001*, volume 2142 of *Lecture Notes in Computer Science*, pages 54–68. Springer-Verlag, September 2001.
http://cs.bath.ac.uk/ag/p/NoncMELLCoS.pdf.

[GS02] Alessio Guglielmi and Lutz Straßburger. A non-commutative extension of MELL. In M. Baaz and A. Voronkov, editors, *LPAR 2002*, volume 2514 of *Lecture Notes in Artificial Intelligence*, pages 231–246. Springer-Verlag, October 2002.
http://www.lix.polytechnique.fr/~lutz/papers/NEL.pdf.

[GS09] Alessio Guglielmi and Lutz Straßburger. A system of interaction and structure V: The exponentials and splitting. Submitted.
http://www.lix.polytechnique.fr/~lutz/papers/NEL-splitting.pdf, 2009.

[GT07] Rajeev Goré and Alwen Tiu. Classical modal display logic in the calculus of structures and minimal cut-free deep inference calculi for S5. *Journal of Logic and Computation*, 17(4):767–794, 2007.
http://users.rsise.anu.edu.au/~tiu/papers/cmdl.pdf.

[Gug03] Alessio Guglielmi. Resolution in the calculus of structures.
http://cs.bath.ac.uk/ag/p/AG10.pdf, 2003.

[Gug04] Alessio Guglielmi. Formalism B.
http://cs.bath.ac.uk/ag/p/AG13.pdf, 2004.

[Gug05] Alessio Guglielmi. The problem of bureaucracy and identity of proofs from the perspective of deep inference. In Paola Bruscoli, François Lamarche, and Charles Stewart, editors, *Structures and Deduction*, pages 53–68. Technische Universität Dresden, 2005. ICALP Workshop. ISSN 1430-211X.
http://cs.bath.ac.uk/ag/p/AG14.pdf.

[Gug07a] Alessio Guglielmi. On the proof complexity of deep inference—Conjecture. Prolog program.
http://cs.bath.ac.uk/ag/p/PrComplDI.plg, 2007.

[Gug07b] Alessio Guglielmi. A system of interaction and structure. *ACM Transactions on Computational Logic*, 8(1):1–64, 2007.
http://cs.bath.ac.uk/ag/p/SystIntStr.pdf.

[Gui06] Yves Guiraud. The three dimensions of proofs. *Annals of Pure and Applied Logic*, 141(1-2):266–295, 2006.
http://www.loria.fr/~guiraudy/recherche/cos1.pdf.

[Jap08] Giorgi Japaridze. Cirquent calculus deepened. *Journal of Logic and Computation*, 18(6):983–1028, 2008.
http://arxiv.org/pdf/0709.1308.

[Jeř09] Emil Jeřábek. Proof complexity of the cut-free calculus of structures. *Journal of Logic and Computation*, 19(2):323–339, 2009.
http://www.math.cas.cz/~jerabek/papers/cos.pdf.

[Kah06a] Ozan Kahramanoğulları. *Nondeterminism and Language Design in Deep Inference*. PhD thesis, Technische Universität Dresden, 2006.
http://www.doc.ic.ac.uk/~ozank/Papers/ozansthesis.pdf.

[Kah06b] Ozan Kahramanoğulları. Reducing nondeterminism in the calculus of structures. In M. Hermann and A. Voronkov, editors, *LPAR 2006*, volume 4246 of *Lecture Notes in Artificial Intelligence*, pages 272–286. Springer-Verlag, 2006.

[Kah07] Ozan Kahramanoğulları. System BV is NP-complete. *Annals of Pure and Applied Logic*, 152(1–3):107–121, 2007.
http://www.doc.ic.ac.uk/~ozank/Papers/bv_npc_apal.pdf.

[Kah08] Ozan Kahramanoğulları. Maude as a platform for designing and implementing deep inference systems. In J. Visser and V. Winter, editors, *Proceedings of the Eighth International Workshop on Rule Based Programming (RULE 2007)*, volume 219 of *Electronic Notes in Theoretical Computer Science*, pages 35–50. Elsevier, 2008.
http://www.doc.ic.ac.uk/~ozank/Papers/rule07.pdf.

[KP89] Jan Krajíček and Pavel Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *Journal of Symbolic Logic*, 54(3):1063–1079, 1989.

[LS05a] François Lamarche and Lutz Straßburger. Constructing free boolean categories. In Prakash Panangaden, editor, *20th Annual IEEE Symposium on Logic in Computer Science*, pages 209–218. IEEE, 2005.
http://www.lix.polytechnique.fr/~lutz/papers/FreeBool-long.pdf.

[LS05b] François Lamarche and Lutz Straßburger. Naming proofs in classical propositional logic. In Paweł Urzyczyn, editor, *Typed Lambda Calculi and Applications*, volume 3461 of *Lecture Notes in Computer Science*, pages 246–261. Springer-Verlag, 2005.
http://www.lix.polytechnique.fr/~lutz/papers/namingproofsCL.pdf.

[LS06] François Lamarche and Lutz Straßburger. From proof nets to the free *-autonomous category. *Logical Methods in Computer Science*, 2(4):3:1–44, 2006.
http://arxiv.org/pdf/cs.LO/0605054.

[Rec76] Robert A. Reckhow. *On the Lengths of Proofs in the Propositional Calculus*. PhD thesis, University of Toronto, 1976.

[SL04] Lutz Straßburger and François Lamarche. On proof nets for multiplicative linear logic with units. In J. Marcinkowski and A. Tarlecki, editors, *CSL 2004*, volume 3210 of *Lecture Notes in Computer Science*, pages 145–159. Springer-Verlag, 2004.
http://www.lix.polytechnique.fr/~lutz/papers/multPN.pdf.

[Sta78] Richard Statman. Bounds for proof-search and speed-up in the predicate calculus. *Annals of Mathematical Logic*, 15:225–287, 1978.

[Sto07] Phiniki Stouppa. A deep inference system for the modal logic S5. *Studia Logica*, 85(2):199–214, 2007.
http://www.iam.unibe.ch/til/publications/pubitems/pdfs/sto07.pdf.

[Str02] Lutz Straßburger. A local system for linear logic. In M. Baaz and A. Voronkov, editors, *LPAR 2002*, volume 2514 of *Lecture Notes in Artificial Intelligence*, pages 388–402. Springer-Verlag, 2002.
http://www.lix.polytechnique.fr/~lutz/papers/lls-lpar.pdf.

[Str03a] Lutz Straßburger. *Linear Logic and Noncommutativity in the Calculus of Structures*. PhD thesis, Technische Universität Dresden, 2003.
http://www.lix.polytechnique.fr/~lutz/papers/dissvonlutz.pdf.

[Str03b] Lutz Straßburger. MELL in the calculus of structures. *Theoretical Computer Science*, 309:213–285, 2003.
http://www.lix.polytechnique.fr/~lutz/papers/els.pdf.

[Tiu06a] Alwen Tiu. A local system for intuitionistic logic. In M. Hermann and A. Voronkov, editors, *LPAR 2006*, volume 4246 of *Lecture Notes in Artificial Intelligence*, pages 242–256. Springer-Verlag, 2006.
http://users.rsise.anu.edu.au/~tiu/localint.pdf.

[Tiu06b] Alwen Tiu. A system of interaction and structure II: The need for deep inference. *Logical Methods in Computer Science*, 2(2:4):1–24, 2006.
http://arxiv.org/pdf/cs.LO/0512036.

[TS96] A.S. Troelstra and H. Schwichtenberg. *Basic Proof Theory*, volume 43 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1996.

UNIVERSITY OF BATH, BATH BA2 7AY, UK,
http://cs.bath.ac.uk/pb/ AND
http://alessio.guglielmi.name/res