Proofs with monotone cuts

Emil Jeřábek*

Institute of Mathematics of the Academy of Sciences

Žitná 25, 115 67 Praha 1, Czech Republic, email: jerabek@math.cas.cz

August 10, 2010

Abstract

Atserias, Galesi, and Pudlák [4] have shown that the monotone sequent calculus *MLK* quasipolynomially simulates proofs of monotone sequents in the full sequent calculus *LK* (or equivalently, in Frege systems). We generalize the simulation to the fragment *MCLK* of *LK* which can prove arbitrary sequents, but restricts cut-formulas to be monotone. We also show that *MLK* as a refutation system for CNFs quasipolynomially simulates *LK*.

1 Introduction

The propositional sequent calculus *LK*, or equivalently, Frege systems, are among the most natural proof systems studied in propositional proof complexity, however the goal of proving superpolynomial lower bounds for these systems remains elusive. The best we can do are lower bounds (even exponential) for some fragments of *LK* obtained by restricting the class of formulas that can appear in a proof, the primary example being bounded-depth Frege systems. Lower bounds on such subsystems of *LK* can make use of structural or combinatorial properties of Boolean functions definable by the class of formulas in question, thus progress in this direction is connected to progress in circuit complexity. (Though this relationship should not be exaggerated: for a well-known example, bounded-depth Frege systems with mod-p counting gates have so far resisted any attempts to prove lower bounds despite that exponential lower bounds on the corresponding circuit class $AC^0[p]$ have been known for about quarter a century.)

One important circuit class for which we have strong lower bounds is the class of monotone circuits. The *monotone sequent calculus MLK*, which is the fragment of *LK* allowing only monotone formulas to appear in any sequent in the proof, was thus introduced with the hope that one could somehow exploit ideas from monotone circuit complexity to show exponential speed-up of *LK* over *MLK*. This intuition turned out to be wrong: Atserias, Galesi, and Pudlák [4] have shown that *MLK* quasipolynomially simulates *LK* (with respect to monotone sequents), moreover, they presented a plausible hypothesis (basically saying that there are

*Supported by grant IAA1019401 of GA AV ČR, grant 1M0545 of MŠMT ČR, and a grant from the John Templeton Foundation.

polynomial-size monotone formulas for threshold functions whose basic properties have polynomial LK-proofs) which implies that the simulation can be improved to fully polynomial. (Some work halfway towards establishing this conjecture has been done in [8].)

A drawback of the usual setup where subsystems of LK are obtained by restricting all formulas in the proof (as described above) is that it also restricts the set of sequents provable in the system. For example, bounded-depth Frege systems can only prove tautologies of bounded depth, and MLK can only prove monotone sequents. An alternative approach is to restrict only *cut-formulas*, and let arbitrary formulas appear in other parts of the proof. This yields a proof system with two desirable properties: on the one hand, it is a complete proof system for full classical propositional logic with no restriction on provable sequents (because it includes the cut-free sequent calculus), on the other hand, it conservatively extends the former approach in that if the end-sequent of a proof as well as all cut-formulas are restricted to a class of formulas (closed under subformulas), then all formulas in the proof are, due to the subformula property.

In this paper, we apply this alternative approach to the monotone calculus. We introduce the proof system $MCLK$, which is identical to LK except that only monotone formulas are admitted as cut-formulas. By the above mentioned conservativity property, $MCLK$ coincides with MLK when proving monotone sequents, in particular the result of Atserias et al. [4] shows that $MCLK$ quasipolynomially simulates LK-proofs of monotone sequents. However, unlike MLK, $MCLK$ can also prove nonmonotone sequents, and a priori it could be much less efficient in this case (the only obvious upper bound is that $MCLK$ includes the cut-free sequent calculus). Our main result shows that this does not happen: $MCLK$ quasipolynomially simulates LK on arbitrary sequents, and if the hypothesis on threshold functions from [4] is true, then the simulation can be made polynomial. As in [4], the idea of the proof is based on slice functions (see Wegener [13]), but we apply them in a different way.

We also consider MLK as a refutation system for CNFs, which is another way of employing a monotone calculus to prove (or rather refute, in this case) nonmonotone formulas. A CNF can be given as a set of clauses, and each clause can be represented as a monotone sequent (in fact: a sequent containing only propositional variables). Then an MLK-refutation is a derivation of the contradictory empty sequent from this set of sequents using the rules of MLK. We extend the simulation to this setup as well: MLK as a refutation system quasipolynomially (potentially polynomially, under the usual hypothesis) simulates LK-refutations of CNFs (or equivalently, Frege proofs of DNFs). In fact, the result is much more general: we show that $MCLK$ quasipolynomially (potentially polynomially) simulates LK-derivations of a sequent from a set of monotone sequents.

2 Preliminaries

We will work with formulas using propositional variables p_i, $i \in \omega$, and the connectives $\wedge, \vee, \neg, \top, \bot$. Such a formula is *monotone* if it does not contain $\neg$. A *sequent* is an expression of the form $\Gamma \vdash \Delta$, where Γ and Δ are finite sets of formulas (thus our sequent calculi will have the structural rules of exchange and contraction for free; we could also dispense with

$$i\ \frac{}{\varphi \vdash \varphi} \qquad cut\ \frac{\Gamma \vdash \varphi, \Delta \qquad \Pi, \varphi \vdash \Lambda}{\Gamma, \Pi \vdash \Delta, \Lambda}$$

$$w\ \frac{\Gamma \vdash \Delta}{\Gamma, \Pi \vdash \Delta, \Lambda}$$

$$\wedge\text{-l}\ \frac{\Gamma, \varphi, \psi \vdash \Delta}{\Gamma, \varphi \wedge \psi \vdash \Delta} \qquad \wedge\text{-r}\ \frac{\Gamma \vdash \varphi, \Delta \qquad \Pi \vdash \psi, \Lambda}{\Gamma, \Pi \vdash \varphi \wedge \psi, \Delta, \Lambda}$$

$$\vee\text{-l}\ \frac{\Gamma, \varphi \vdash \Delta \qquad \Pi, \psi \vdash \Lambda}{\Gamma, \Pi, \varphi \vee \psi \vdash \Delta, \Lambda} \qquad \vee\text{-r}\ \frac{\Gamma \vdash \varphi, \psi, \Delta}{\Gamma \vdash \varphi \vee \psi, \Delta}$$

$$\neg\text{-l}\ \frac{\Gamma \vdash \varphi, \Delta}{\Gamma, \neg\varphi \vdash \Delta} \qquad \neg\text{-r}\ \frac{\Gamma, \varphi \vdash \Delta}{\Gamma \vdash \neg\varphi, \Delta}$$

$$\bot\text{-l}\ \frac{}{\bot \vdash} \qquad \top\text{-r}\ \frac{}{\vdash \top}$$

Table 1: Rules of the sequent calculus

weakening, but we decided to keep it for convenience). An *LK-proof* of a sequent $\Gamma \vdash \Delta$ is a finite sequence $\Gamma_0 \vdash \Delta_0, \dots, \Gamma_k \vdash \Delta_k$ of sequents such that $(\Gamma_k \vdash \Delta_k) = (\Gamma \vdash \Delta)$, and each $\Gamma_i \vdash \Delta_i$ is derived from some of the sequents $\{\Gamma_j \vdash \Delta_j \mid j < i\}$ by an *LK*-rule, as listed in Table 1. The *number of lines* in the proof is $k+1$, and its *size* is the total number of symbols. An *LK*-proof of a formula φ is a proof of the sequent $\vdash \varphi$.

A sequent is monotone, if all of its formulas are. An *MLK-proof* is an *LK*-proof consisting only of monotone sequents (in particular, no $\neg$-l or $\neg$-r rules can be used). We introduce our proof system *MCLK* as follows: an *MCLK-proof* is an *LK*-proof where in every instance of the cut rule, the formula φ is monotone. The basic relationship of *MLK* to *MCLK* is given by the following observation:

Proposition 2.1 *MLK-proofs are exactly MCLK-proofs of monotone sequents.*

Proof: Clearly, an *MLK*-proof is also an *MCLK*-proof, and its end-sequent is monotone. Conversely, consider an *MCLK*-proof of a monotone sequent. By the subformula property, every formula in the proof is a subformula of some formula in the end-sequent or of a cut-formula, and as such it is monotone. Thus the proof is an *MLK*-proof. □

The following result is known about the complexity of *MLK*:

Theorem 2.2 (Atserias et al. [4]) *MLK quasipolynomially simulates LK-proofs of monotone sequents. More precisely, if a monotone sequent in n variables has an LK-proof of size s, it also has an MLK-proof of size $n^{O(\log n)} s^{O(1)}$ with $s^{O(1)}$ lines.* □

We do not formally introduce the concept of quasipolynomial simulation, as we will always give the size of the resulting proof explicitly as here (the reason being that it is better than a general quasipolynomial bound, which is $2^{(\log s)^{O(1)}}$). However, we note that here and below, all simulations are actually constructive in the sense that given an LK-proof, we can find the corresponding MLK-proof in time polynomial in the size of the output (i.e., $n^{O(\log n)}s^{O(1)}$).

The main ingredient in the simulation in [4] are *threshold functions*, i.e., Boolean functions $\theta^n_k \colon 2^n \to 2$ defined by

$$\theta^n_k(x_0, \dots, x_{n-1}) = 1 \Leftrightarrow \bigl|\{i < n \mid x_i = 1\}\bigr| \geq k.$$

Using a divide-and-conquer approach

$$T^n_k(p_0, \dots, p_{n-1}) = \bigvee_{\substack{i \leq k \\ i \leq n/2}} \Bigl(T^{\lfloor n/2 \rfloor}_i(p_0, \dots, p_{\lfloor n/2 \rfloor - 1}) \wedge T^{\lceil n/2 \rceil}_{k-i}(p_{\lfloor n/2 \rfloor}, \dots, p_{n-1})\Bigr),$$

we can define monotone formulas T^n_k computing θ^n_k of size $n^{O(\log n)}$. The main properties of these formulas have short proofs:

Theorem 2.3 (Atserias et al. [3]) *The sequents*

$$\vdash T^n_0(p_0, \dots, p_{n-1}) \tag{1}$$

$$T^n_{n+1}(p_0, \dots, p_{n-1}) \vdash \tag{2}$$

$$T^n_k(p_0, \dots, p_{i-1}, \bot, p_{i+1}, \dots, p_{n-1}) \vdash T^n_{k+1}(p_0, \dots, p_{i-1}, \top, p_{i+1}, \dots, p_{n-1}) \tag{3}$$

have MLK-proofs of size $n^{O(\log n)}$ with $n^{O(1)}$ lines. □

Using carry-save addition, it is possible to construct polynomial-size (nonmonotone) formulas for the threshold functions such that (1)–(3) have polynomial-size LK-proofs (Buss [5]). Actually, there exist polynomial-size *monotone* formulas for θ^n_k, however the known constructions are either randomized (Valiant [12]) or rather complicated (sorting networks by Ajtai, Komlós, and Szemerédi [1, 2]), thus their basic properties are not known to be shortly provable.

Theorem 2.4 ([4]) *Assume that there are monotone formulas such that* (1)–(3) *have polynomial-size LK-proofs. Then MLK polynomially simulates LK-proofs of monotone sequents. I.e., if a monotone sequent has an LK-proof of size s, then it has an MLK-proof of size $s^{O(1)}$.* □

Note in particular that the result only asks for proofs of (1)–(3) in LK, rather than MLK. Some progress towards establishing the hypothesis of Theorem 2.4 has been achieved in [8] by formalizing a variant of the Ajtai–Komlós–Szemerédi network in a suitable theory of bounded arithmetic corresponding to LK; the missing piece is formalization of an expander graph construction.

3 Elimination of constants

Before we get to the main results, we will clarify one issue about the definition of the monotone calculi, namely the role of the truth constants $\top, \bot$. For convenience, we defined MLK and

MCLK so that both $\top, \bot$ can appear in monotone formulas; however, it is just as reasonable to restrict monotone formulas to $\wedge$ and $\vee$ only. We will show that this choice does not matter: we can eliminate $\top, \bot$ from a proof with no increase in size if they do not appear in the end-sequent. This is an expected result, nevertheless we feel that it is better to state and prove it explicitly, so that we do not have to worry about it later.

Theorem 3.1 *If a sequent with no occurrence of $\top$ or $\bot$ has an MCLK-proof of size s, it also has an MCLK-proof of size at most s with no occurrence of $\top$ or $\bot$.*

Proof: For each formula φ, let us define a formula $\overline{\varphi}$ by induction on the complexity of φ:

$$\overline{\varphi} = \varphi \text{ if } \varphi \text{ is a variable or } \top \text{ or } \bot, \qquad \overline{\neg\varphi} = \begin{cases} \top & \text{if } \overline{\varphi} = \bot, \\ \bot & \text{if } \overline{\varphi} = \top, \\ \neg\overline{\varphi} & \text{otherwise,} \end{cases}$$

$$\overline{\varphi \wedge \psi} = \begin{cases} \bot & \text{if } \overline{\varphi} = \bot \text{ or } \overline{\psi} = \bot, \\ \overline{\psi} & \text{if } \overline{\varphi} = \top, \\ \overline{\varphi} & \text{if } \overline{\psi} = \top, \\ \overline{\varphi} \wedge \overline{\psi} & \text{otherwise,} \end{cases} \qquad \overline{\varphi \vee \psi} = \begin{cases} \top & \text{if } \overline{\varphi} = \top \text{ or } \overline{\psi} = \top, \\ \overline{\psi} & \text{if } \overline{\varphi} = \bot, \\ \overline{\varphi} & \text{if } \overline{\psi} = \bot, \\ \overline{\varphi} \vee \overline{\psi} & \text{otherwise.} \end{cases}$$

That is, we propagate constants upwards until they either disappear, or we end up with evaluation of the whole formula. Observe that $|\overline{\varphi}| \leq |\varphi|$, $\overline{\varphi}$ is monotone if φ is, $\overline{\varphi}$ contains no occurrence of $\top$ or $\bot$ unless it *is* $\top$ or $\bot$, and $\overline{\varphi} = \varphi$ if φ contains no constant. Let π be an *MCLK*-proof whose end-sequent contains no constant. We construct π' by applying the following transformation to π:

(i) we replace every formula φ with its translation $\overline{\varphi}$,

(ii) we remove all occurrences of $\top$ from antecedents and $\bot$ from succedents,

(iii) we delete all sequents which contain $\bot$ in antecedent or $\top$ in succedent.

We obtain a sequence of sequents with no occurrence of $\top$ or $\bot$, which ends with the original end-sequent of π. We claim that the sequence is a valid *MCLK*-proof. For example, assume that a sequent was derived in π by an application

$$(*)\ \frac{\Gamma \vdash \varphi, \Delta \qquad \Pi \vdash \psi, \Lambda}{\Gamma, \Pi \vdash \varphi \wedge \psi, \Delta, \Lambda}$$

of the $\wedge$-r rule, and that the corresponding sequent was not deleted in step (iii). This means that there is no $\bot$ in $\overline{\Gamma} \cup \overline{\Pi}$, no $\top$ in $\overline{\Delta} \cup \overline{\Lambda}$, and at least one of $\overline{\varphi}$ and $\overline{\psi}$ is different from $\top$, therefore at least one sequent corresponding to the two assumptions of the rule is retained in π'. If, say, $\overline{\varphi}$ is $\bot$, then also $\overline{\varphi \wedge \psi} = \bot$, and $(*)$ translates to an instance

$$\frac{\overline{\Gamma}' \vdash \overline{\Delta}'}{\overline{\Gamma}', \overline{\Pi}' \vdash \overline{\Delta}', \overline{\Lambda}'}$$

of weakening, where primes denote the removal of constants in step (ii). We may thus assume that neither $\overline{\varphi}$ nor $\overline{\psi}$ is $\bot$. If $\overline{\varphi}$ is $\top$, then $(*)$ translates to another instance

$$\frac{\overline{\Pi}' \vdash \overline{\psi}, \overline{\Lambda}'}{\overline{\Gamma}', \overline{\Pi}' \vdash \overline{\psi}, \overline{\Delta}', \overline{\Lambda}'}$$

of weakening. Similarly if $\overline{\psi} = \top$. Finally, if neither $\overline{\varphi}$ nor $\overline{\psi}$ is a constant, then $(*)$ translates to an instance

$$\frac{\overline{\Gamma}' \vdash \overline{\varphi}, \overline{\Delta}' \qquad \overline{\Pi}' \vdash \overline{\psi}, \overline{\Lambda}'}{\overline{\Gamma}', \overline{\Pi}' \vdash \overline{\varphi} \wedge \overline{\psi}, \overline{\Delta}', \overline{\Lambda}'}$$

of $\wedge$-r. The other rules are handled in a similar way, we leave the details to the reader. □

Corollary 3.2 *If a sequent with no occurrence of $\top$ or $\bot$ has an MLK-proof of size s, it also has an MLK-proof of size at most s with no occurrence of $\top$ or $\bot$.*

Proof: By Proposition 2.1 and Theorem 3.1. □

Note that the proof of Theorem 3.1 also applies to *LK*, but there we could obtain it trivially (albeit with a linear size increase) by replacing $\bot, \top$ with $p \wedge \neg p, p \vee \neg p$ (respectively).

4 Simulation of *LK* by monotone cuts

In this section, we prove our main result on simulation of *LK* by *MCLK*:

Theorem 4.1 *MCLK quasipolynomially simulates LK: if a sequent in n variables has an LK-proof of size s, it also has an MCLK-proof of size $n^{O(\log n)} s^{O(1)}$ with $s^{O(1)}$ lines. If the assumption of Theorem 2.4 holds true, then it has an MCLK-proof of size $s^{O(1)}$.*

Let T^n_k denote either the formulas of size $n^{O(\log n)}$ employed in Theorem 2.3, or polynomial-size formulas for threshold functions given by the assumption of Theorem 2.4. Let us also fix $k \leq n + 1$.

We will use the following consequence of Theorem 2.3:

Lemma 4.2 ([4]) *The sequents*

$$T^n_k(p_0, \ldots, p_{n-1}) \vdash p_i, T^n_k(p_0, \ldots, p_{i-1}, \bot, p_{i+1}, \ldots, p_{n-1}), \tag{4}$$

$$T^n_k(p_0, \ldots, p_{i-1}, \bot, p_{i+1}, \ldots, p_{n-1}), p_i \vdash T^n_{k+1}(p_0, \ldots, p_{n-1}) \tag{5}$$

have MLK-proofs of size $n^{O(\log n)}$ with $n^{O(1)}$ lines for every $i < n$. If the assumption of Theorem 2.4 holds true, they have MLK-proofs of size $n^{O(1)}$. □

For every formula $\varphi(p_0, \ldots, p_{n-1})$, we define a monotone formula $\overline{\varphi}(p_0, \ldots, p_{n-1})$ as follows: we use De Morgan rules to push all negations down to variables, and then we replace each $\neg p_i$ with the formula

$$T^n_k(p_0, \ldots, p_{i-1}, \bot, p_{i+1}, \ldots, p_{n-1}).$$

Lemma 4.3 *For any formula $\varphi(p_0, \ldots, p_{n-1})$, the sequents*

$$T^n_k(p_0, \ldots, p_{n-1}) \vdash \overline{\varphi}, \overline{\neg\varphi}, \tag{6}$$

$$\overline{\varphi}, \overline{\neg\varphi} \vdash T^n_{k+1}(p_0, \ldots, p_{n-1}), \tag{7}$$

$$T^n_k(p_0, \ldots, p_{n-1}), \varphi \vdash \overline{\varphi}, \tag{8}$$

$$\overline{\varphi} \vdash \varphi, T^n_{k+1}(p_0, \ldots, p_{n-1}) \tag{9}$$

have MCLK-proofs of size $n^{O(\log n)}|\varphi|^{O(1)}$ with $(n|\varphi|)^{O(1)}$ lines. If the assumption of Theorem 2.4 *holds true, they have MCLK-proofs of size $(n|\varphi|)^{O(1)}$.*

Proof: (6): By induction on complexity of φ. If φ is a variable, the sequent is a restatement of (4). If φ is $\bot$ or $\top$, then $T^n_k \vdash \top, \bot$ is an instance of $\top$-r. The induction step for $\neg$ is trivial as $\overline{\neg\neg\varphi} = \overline{\varphi}$. The induction step for $\wedge$ proceeds by

$$\wedge\text{-r}\,\dfrac{T^n_k(p_0, \ldots, p_{n-1}) \vdash \overline{\varphi}, \overline{\neg\varphi} \qquad T^n_k(p_0, \ldots, p_{n-1}) \vdash \overline{\psi}, \overline{\neg\psi}}{\vee\text{-r}\,\dfrac{T^n_k(p_0, \ldots, p_{n-1}) \vdash \overline{\varphi} \wedge \overline{\psi}, \overline{\neg\varphi}, \overline{\neg\psi}}{T^n_k(p_0, \ldots, p_{n-1}) \vdash \overline{\varphi} \wedge \overline{\psi}, \overline{\neg\varphi} \vee \overline{\neg\psi}}}$$

where $\overline{\varphi \wedge \psi} = \overline{\varphi} \wedge \overline{\psi}$ and $\overline{\neg(\varphi \wedge \psi)} = \overline{\neg\varphi} \vee \overline{\neg\psi}$. The induction step for $\vee$ is similar.

(7): The proof is analogous to (6), except that we use (5) instead of (4).

(8): We prove the sequents

$$T^n_k(p_0, \ldots, p_{n-1}), \varphi \vdash \overline{\varphi},$$
$$T^n_k(p_0, \ldots, p_{n-1}) \vdash \varphi, \overline{\neg\varphi}$$

simultaneously by induction on complexity of φ. We use (4) when φ is atomic. The induction step for $\neg$ goes by

$$\neg\text{-l}\,\dfrac{T^n_k(p_0, \ldots, p_{n-1}) \vdash \varphi, \overline{\neg\varphi}}{T^n_k(p_0, \ldots, p_{n-1}), \neg\varphi \vdash \overline{\neg\varphi}} \qquad \neg\text{-r}\,\dfrac{T^n_k(p_0, \ldots, p_{n-1}), \varphi \vdash \overline{\varphi}}{T^n_k(p_0, \ldots, p_{n-1}) \vdash \neg\varphi, \overline{\varphi}}$$

and the induction steps for $\wedge$ and $\vee$ are similar to the proof of (6).

The proof of (9) is analogous. □

Proof (*of Theorem* 4.1)*:* Let π be an *LK*-proof of a sequent $\Theta \vdash \Xi$. We fix $k \le n+1$, and we consider the corresponding translation $\overline{\varphi}$ as defined above. We construct $\overline{\pi}$ by replacing each sequent $\Gamma \vdash \Delta$ in π with

$$T^n_k(p_0, \ldots, p_{n-1}), \overline{\Gamma} \vdash \overline{\Delta}, T^n_{k+1}(p_0, \ldots, p_{n-1}),$$

where $\overline{\Gamma} = \{\overline{\varphi} \mid \varphi \in \Gamma\}$. If a sequent was derived in π by one of the rules i, w, cut, $\bot$-l, or $\top$-r, then its translation is derived by an instance of the same rule in $\overline{\pi}$. In fact, the same also holds for $\wedge$-l, $\wedge$-r, $\vee$-l, and $\vee$-r, because $\overline{\varphi \wedge \psi} = \overline{\varphi} \wedge \overline{\psi}$ and $\overline{\varphi \vee \psi} = \overline{\varphi} \vee \overline{\psi}$. If a sequent was derived in π by an instance

$$\dfrac{\Gamma, \varphi \vdash \Delta}{\Gamma \vdash \neg\varphi, \Delta}$$

of $\neg$-r, we derive its translation in $\overline{\pi}$ by a (monotone) cut

$$\frac{T^n_k(p_0,\dots,p_{n-1}),\overline{\Gamma},\overline{\varphi}\vdash\overline{\Delta},T^n_{k+1}(p_0,\dots,p_{n-1})\qquad T^n_k(p_0,\dots,p_{n-1})\vdash\overline{\varphi},\overline{\neg\varphi}}{T^n_k(p_0,\dots,p_{n-1}),\overline{\Gamma}\vdash\overline{\neg\varphi},\overline{\Delta},T^n_{k+1}(p_0,\dots,p_{n-1})}$$

with the sequent (6) from Lemma 4.3. Instances of $\neg$-l are handled similarly, using (7).

In this way, we obtain a valid *MLK*-proof of the sequent

$$T^n_k(p_0,\dots,p_{n-1}),\overline{\Theta}\vdash\overline{\Xi},T^n_{k+1}(p_0,\dots,p_{n-1}).$$

For every formula $\varphi\in\Theta$, we use a cut with (8) to replace $\overline{\varphi}$ with φ. (Note that the cut formula, $\overline{\varphi}$, is monotone.) Similarly, we replace each $\overline{\psi}\in\overline{\Xi}$ with ψ by a cut with (9). We obtain an *MCLK*-proof of

$$T^n_k(p_0,\dots,p_{n-1}),\Theta\vdash\Xi,T^n_{k+1}(p_0,\dots,p_{n-1}).$$

We do this for every $k\le n+1$, and we join all these proofs together by $n+1$ cuts to get an *MCLK*-proof of

$$T^n_0(p_0,\dots,p_{n-1}),\Theta\vdash\Xi,T^n_{n+1}(p_0,\dots,p_{n-1}).$$

Finally, using cuts with (1) and (2), we get an *MCLK*-proof of

$$\Theta\vdash\Xi.$$

□

5 *MLK* as a refutation system

We have defined *MCLK* with the intention of creating a proof system which would share the main properties of *MLK*, but could be used to derive nonmonotone formulas. There is another way of achieving this, namely to use *MLK* as a *refutation system* for CNFs. A CNF-formula can be presented as a set of clauses

$$p_{i_1}\vee p_{i_2}\vee\cdots\vee p_{i_k}\vee\neg p_{j_1}\vee\neg p_{j_2}\vee\cdots\vee\neg p_{j_l}.$$

Such a clause C can in turn be identified with the monotone sequent $C^{\vdash}$ defined as

$$p_{j_1},p_{j_2},\dots,p_{j_l}\vdash p_{i_1},p_{i_2},\dots,p_{i_k}$$

consisting only of propositional variables.

Definition 5.1 Let S be a set of clauses. An *MLK-refutation* of S is a finite sequence $\{\Gamma_i\vdash\Delta_i\mid i\le k\}$ of sequents such that $\Gamma_k\vdash\Delta_k$ is the contradictory empty sequent $\vdash$, and each $\Gamma_i\vdash\Delta_i$ is either $C^{\vdash}$ for an element C of S, or is derived from some of the sequents $\{\Gamma_j\vdash\Delta_j\mid j<i\}$ by an *MLK*-rule. *LK*-refutations are defined similarly.

Notice that *MLK* is complete as a refutation system for CNFs, as it includes resolution (indeed, resolution is the fragment of *MLK* which only allows cut as a rule of inference).

Notice also that it makes no difference whether we formulate the refutation system using *MLK*-rules or *MCLK*-rules: no nonmonotone formulas can sneak in because of the subformula property. We can, however, generalize the setup to derivations of arbitrary sequents from sets of sequents, and then it is meaningful to consider *MCLK*-rules to allow for nonmonotone sequents.

Definition 5.2 Let S be a set of sequents. An *MCLK-derivation* of a sequent $\Gamma \vdash \Delta$ from S is a finite sequence of sequents ending with $\Gamma \vdash \Delta$ such that every element of the sequence is derived by an *MCLK*-rule from previous ones or is a member of S. *LK*-derivations and *MLK*-derivations are defined similarly.

Observe that *MCLK*-derivations of *monotone* sequents from S are in fact *MLK*-derivations, due to the subformula property (and in particular, they cannot make use of any nonmonotone sequents in S).

Unrestricted *LK*-derivations have the same power as *LK* employed in the usual way as a proof system:

Definition 5.3 The *characteristic formula* $(\Gamma \vdash \Delta)^{\rightarrow}$ of a sequent $\Gamma \vdash \Delta$ is the formula

$$\bigvee_{\varphi \in \Gamma} \neg\varphi \vee \bigvee_{\psi \in \Delta} \psi$$

(with arbitrary bracketing of the disjunctions).

Proposition 5.4 (folklore) *The following are constructible from each other in polynomial time:*

(*i*) *an LK-derivation of a sequent A from a finite set S of sequents,*

(*ii*) *an LK-proof of the formula*

$$\bigvee_{B \in S} \neg B^{\rightarrow} \vee A^{\rightarrow}.$$

(*Either the input of the algorithm includes S, or the disjunction in* (*ii*) *should be restricted to those B actually used in the proof.*)

Proof: (i) $\rightarrow$ (ii): we weaken all sequents in the derivation by including $\bigvee_{B\in S} \neg B^{\rightarrow}$ on the right-hand side. For each initial sequent $(\Pi \vdash \Lambda) \in S$ used in the derivation, we include the subproof

$$\dfrac{\dfrac{\cdots \quad \varphi \vdash \varphi \quad (\varphi \in \Lambda) \quad \cdots \quad \dfrac{\varphi \vdash \varphi}{\neg\varphi, \varphi \vdash}\,\neg\text{-l} \quad (\varphi \in \Pi) \quad \cdots}{(\Pi \vdash \Lambda)^{\rightarrow}, \Pi \vdash \Lambda}\,\vee\text{-l}}{\Pi \vdash \Lambda, \bigvee_{B \in S} \neg B^{\rightarrow}}\,\neg\text{-r,w,}\vee\text{-r.}$$

Writing $A = (\Gamma \vdash \Delta)$, we use $\neg$-r and $\vee$-r to derive the end-sequent $\vdash \bigvee_{B\in S} \neg B^{\rightarrow} \vee (\Gamma \vdash \Delta)^{\rightarrow}$ from $\Gamma \vdash \Delta, \bigvee_{B\in S} \neg B^{\rightarrow}$.

(ii) $\to$ (i): if $A = (\Gamma \vdash \Delta)$, we construct a proof of $A^{\to}, \Gamma \vdash \Delta$ as above. Similarly, for each $B = (\Pi \vdash \Lambda) \in S$, we construct a derivation of $\neg B^{\to} \vdash$ from B using $\neg$-r, $\vee$-r, and $\neg$-l. We infer

$$\bigvee_{B \in S} \neg B^{\to} \vee A^{\to}, \Gamma \vdash \Delta$$

by $\vee$-l, and $\Gamma \vdash \Delta$ by a cut. $\square$

Corollary 5.5 *The following are constructible from each other in polynomial time:*

(i) an LK-refutation of a set $C = \{C_i \mid i < t\}$ *of clauses,*

(ii) an LK-proof of the formula $\neg \bigwedge_{i<t} C_i$. $\square$

The main observation of this section is that the simulation of *LK* by *MCLK* generalizes to derivations from sets of monotone sequents.

Theorem 5.6 *If a sequent A has an LK-derivation of size s from a set S of monotone sequents, where $S \cup \{A\}$ use n variables, then it also has an MCLK-derivation from S of size $n^{O(\log n)} s^{O(1)}$ with $s^{O(1)}$ lines. If the assumption of Theorem* 2.4 *holds true, then it has an MCLK-derivation from S of size $s^{O(1)}$.*

Proof: We follow the same proof as for Theorem 4.1. We only need to derive translations

$$T_k^n(p_0, \ldots, p_{n-1}), \overline{\Gamma} \vdash \overline{\Delta}, T_{k+1}^n(p_0, \ldots, p_{n-1})$$

of initial sequents $\Gamma \vdash \Delta$ from S. However, this is trivial: $\overline{\Gamma} = \Gamma$ and $\overline{\Delta} = \Delta$ since the sequent is monotone, thus we can derive the translation by weakening from $\Gamma \vdash \Delta$ itself. $\square$

Corollary 5.7 *If a CNF in n variables has an LK-refutation of size s, it has an MLK-refutation of size $n^{O(\log n)} s^{O(1)}$ with $s^{O(1)}$ lines. If the assumption of Theorem* 2.4 *holds true, then it has an MLK-refutation of size $s^{O(1)}$.* $\square$

We remark that the proof of Theorem 3.1 (elimination of constants) also generalizes to *MCLK*-derivations from sets of constant-free sequents.

The restriction to S consisting of monotone sequents only in Theorem 5.6 is necessary, because otherwise an *MCLK*-derivation of A from S might not exist at all. For example, *LK* derives $A = (\ \vdash\ )$ from $S = \{\vdash p, \vdash \neg p\}$ by

$$\dfrac{\vdash \neg p \qquad \dfrac{\vdash p}{\neg p \vdash}\,\neg\text{-l}}{\vdash}\,\text{cut},$$

but there is no *MCLK*-derivation of A from S: by the subformula property, every formula in the derivation is a subformula of a formula in A (there are none) or of a cut-formula, and therefore it is monotone; in particular, the axiom $\vdash \neg p$ cannot actually appear in the derivation, and of course there is no derivation of $\vdash$ from $\vdash p$ alone. In general, nonmonotone parts of sequents from S used in an $MCLK$-derivation have to be reflected in the end-sequent, they cannot disappear.

Even when there *is* an $MCLK$-derivation of A from S, we cannot in general expect its size to be related to the size of an LK-derivation. For example, let S be as above, and let A be an arbitrary tautological sequent not involving p. There is a trivial polynomial-size LK-derivation of A from S (namely, the one above extended by weakening). On the other hand, an $MCLK$-derivation of A from S (which exists as A is tautological by itself) cannot involve the axiom $\vdash \neg p$, and then by substituting $\top$ for p we see that without loss of generality it does not use the axiom $\vdash p$ either. Thus, in general A cannot have a (quasi)polynomial-size $MCLK$-derivation from S unless $MCLK$ is (quasi)polynomially bounded.

Leaving proof complexity aside, the reader may wonder whether there is a criterion determining when an $MCLK$-derivation (of arbitrary size) of A from S exists. (For LK, the answer is obvious: if and only if A follows from S in classical logic.) We can obtain one by a straightforward adaptation of the 3-valued semantics of the cut-free sequent calculus (see Schütte [11], Girard [6]).

Definition 5.8 A *Schütte valuation* is a mapping v from the set of propositional formulas to $\{0, 1, *\}$ (where $*$ stands for "undefined") satisfying the following conditions for all formulas φ, ψ:

- if $v(\neg\varphi) = 0$, then $v(\varphi) = 1$,
- if $v(\neg\varphi) = 1$, then $v(\varphi) = 0$,
- if $v(\varphi \wedge \psi) = 1$, then $v(\varphi) = v(\psi) = 1$,
- if $v(\varphi \wedge \psi) = 0$, then $v(\varphi) = 0$ or $v(\psi) = 0$,
- if $v(\varphi \vee \psi) = 1$, then $v(\varphi) = 1$ or $v(\psi) = 1$,
- if $v(\varphi \vee \psi) = 0$, then $v(\varphi) = v(\psi) = 0$,
- $v(\top) \neq 0$,
- $v(\bot) \neq 1$.

A valuation v *satisfies* a sequent $\Gamma \vdash \Delta$ unless $v(\varphi) = 1$ for all $\varphi \in \Gamma$ and $v(\varphi) = 0$ for all $\varphi \in \Delta$. A valuation v is *monotonely full* if $v(\varphi) \neq *$ for every monotone formula φ.

It is easy to see that any Schütte valuation can be extended to a usual Boolean valuation, hence we can think about it as follows: we take a Boolean valuation, and restrict it to a subset of formulas which includes all monotone formulas, and includes suitable witnesses ensuring that (defined) values of compound formulas are computed correctly. Notice also that Schütte semantics is not truth-functional, the value of a compound formula may not be uniquely determined by values of its components: for example, if $v(\varphi) = v(\psi) = 1$, then $v(\varphi \wedge \psi)$ can be either 1 or $*$.

Proposition 5.9 *Let A be a sequent and S a set of sequents. The following are equivalent:*

(*i*) *There exists an MCLK-derivation of A from S.*

(*ii*) *For every monotonely full Schütte valuation v, if v satisfies all sequents in S, then it satisfies A.*

Proof (*sketch*)*:* (i) → (ii): by induction on the length of the derivation. *LK*-rules other than cut are sound for all Schütte valuations. Monotone cuts are sound because the valuation is monotonely full.

(ii) → (i): Assume that A is not *MCLK*-derivable from S. Consider pairs $X \vdash Y$ of (not necessarily finite) sets of formulas, ordered by product of inclusions (i.e., $(X \vdash Y) \leq (X' \vdash Y')$ iff $X \subseteq X'$ and $Y \subseteq Y'$). By Zorn's lemma, there exists a maximal $(X \vdash Y) \geq A$ such that no finite $B \leq (X \vdash Y)$ is *MCLK*-derivable from S. As $X \cap Y = \varnothing$ due to the identity axiom, we can define

$$v(\varphi) = \begin{cases} 1, & \text{if } \varphi \in X, \\ 0, & \text{if } \varphi \in Y, \\ *, & \text{otherwise.} \end{cases}$$

Then it is easy to see that v is a monotonely full Schütte valuation which satisfies S, but refutes A. □

Remark 5.10 The proof did not use any special property of monotone formulas, it actually gives the following more general result: if Φ is an arbitrary set of formulas, then *LK*-derivations with cut-formulas restricted to Φ are sound and complete with respect to Φ-full Schütte valuations.

Example 5.11 There is no *MCLK*-derivation of $p \vee \neg q \vdash \neg p$ from $S = \{q \vdash \neg p, \neg q \vdash \neg p\}$ (even though its existence does not seem to obviously violate the subformula property). To see this, let w be the Boolean valuation such that $w(p) = 1$ and $w(q) = 0$, and let v be the restriction of w defined only for monotone formulas, $\neg p$, and $p \vee \neg q$. Then v is a monotonely full Schütte valuation satisfying S and refuting $p \vee \neg q \vdash \neg p$.

6 Variants and problems

So far we have only worked with sequence-like (or dag-like) proofs and derivations. Alternatively, we may consider *tree-like* derivations, where each sequent is used at most once as a premise of a rule of inference. It is well-known that tree-like *LK* is polynomially equivalent to *LK* (Krajíček [9, 10]), but since the proof of this result heavily relies on implications and cuts, it is not known whether it holds for *MLK* or *MCLK* as well. However, the transformation used in the simulation of *LK* by *MCLK* preserves tree-likeness, provided that we start with tree-like proofs of the basic properties of our threshold formulas:

Theorem 6.1 *If a sequent A has an LK-derivation of size s from a set S of monotone sequents, where $S \cup \{A\}$ use n variables, then it also has a tree-like MCLK-derivation from S of size $n^{O(\log n)} s^{O(1)}$.* □

We do not know whether we can make the number of lines in the tree-like *MCLK*-derivation polynomial[1] as in Theorem 5.6. Also, the assumption of Theorem 2.4 does not seem to suffice to produce polynomial-size tree-like *MCLK*-derivations: the *LK*-proofs in the assumption can be assumed to be tree-like, but the inductive construction of their *MLK*-counterparts in [4] does not preserve tree-likeness. We would need to strengthen the assumption to say that there are polynomial-size tree-like *MLK*-proofs of the sequents, and we have so far no evidence supporting this stronger assumption.

An important strengthening of *LK* is the *extension Frege* (*EF*) system, which allows on top of *LK* to use abbreviations through introduction of *extension axioms* of the form

$$q \leftrightarrow \varphi,$$

where q is a variable not contained in φ, the conclusion of the proof, or the previous part of the proof. The actual extension axioms are not well suited for the context of monotone sequent calculus as $\leftrightarrow$ is a nonmonotone connective, however it is well-known that *EF* is equivalent to a modification of *LK* which operates with Boolean *circuits* instead of formulas, and in this formulation it makes perfect sense for any set of connectives, even independently of their meaning in a particular logic.

For definiteness, we will follow the approach presented in Jeřábek [7]. Two circuits are *similar* if they unwind to the same formula (note that this condition is checkable in polynomial time, in fact, even in *NL*). The circuit-*LK* proof system operates with sequents composed of circuits. It has a structural derivation rule which allows to replace any circuit in a sequent with a similar circuit, and it includes the usual rules of *LK* (note that expressions like $\varphi \wedge \psi$ in the definition of *LK*-rules are ambiguous in principle when applied to circuits, but the similarity rule ensures that all their possible readings are derivable from each other and therefore no confusion can arise). Alternatively, we could replace (with a polynomial blow-up in size) the similarity rule by its special case where one can only merge or unmerge two nodes with identical inputs and labelled by the same connective.

The proof of Theorem 5.6 goes through for circuits as well as for formulas. Moreover, the divide-and-conquer monotone formulas for threshold functions are easily seen to be implementable by polynomial-size monotone circuits, and their basic properties have polynomial-size circuit-*MLK* proofs. We obtain:

Theorem 6.2 *If a sequent A has a circuit-LK-derivation of size s from a set of monotone sequents S, then it has a circuit-MCLK-derivation from S of size $s^{O(1)}$.* □

We do not know whether we can make the circuit-*MCLK*-derivation tree-like. Indeed, by the standard argument (see e.g. [10]) the size of a circuit proof is polynomially related to the number of lines in a proof with formulas, thus this boils down to the above mentioned problem whether tree-like *MCLK* can simulate *LK* in polynomially many lines.

[1]Atserias et al. [4] claim that there are tree-like proofs of Theorem 2.3 and Lemma 4.2 as well as the main simulation (Theorem 2.2) with polynomially many lines, but this appears to be an error. For example, already their Lemma 2 (taken in turn from [3], where neither tree-likeness nor line count is explicitly mentioned) applied to a quasipolynomial-size formula requires quasipolynomially many lines if the proof is to be tree-like, and the same goes for their Lemma 3.

Finally, since we already mentioned several open problems in this section, let us restate here for the record the main problem: can the quasipolynomial simulations in Theorems 2.2, 4.1, and 5.6 be made polynomial, and specifically, is the assumption of Theorem 2.4 true?

7 Acknowledgement

I would like to thank Pavel Pudlák for useful suggestions and enlightening comments.

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi, *Sorting in $c \log n$ parallel steps*, Combinatorica 3 (1983), no. 1, pp. 1–19.

[2] ———, *An $O(n \log n)$ sorting network*, in: Proceedings of the 15th Annual ACM Symposium on Theory of Computing, 1983, pp. 1–9.

[3] Albert Atserias, Nicola Galesi, and Ricard Gavaldà, *Monotone proofs of the Pigeon Hole Principle*, Mathematical Logic Quarterly 47 (2001), no. 4, pp. 461–474.

[4] Albert Atserias, Nicola Galesi, and Pavel Pudlák, *Monotone simulations of non-monotone proofs*, Journal of Computer and System Sciences 65 (2002), no. 4, pp. 626–638.

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

[6] Jean-Yves Girard, *Proof theory and logical complexity*, vol. I, Bibliopolis, Naples, 1987.

[7] Emil Jeřábek, *Dual weak pigeonhole principle, Boolean complexity, and derandomization*, Annals of Pure and Applied Logic 129 (2004), pp. 1–37.

[8] ———, *A sorting network in bounded arithmetic*, submitted, 2008.

[9] Jan Krajíček, *Lower bounds to the size of constant-depth propositional proofs*, Journal of Symbolic Logic 59 (1994), no. 1, pp. 73–86.

[10] ———, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications vol. 60, Cambridge University Press, 1995.

[11] Kurt Schütte, *Syntactical and semantical properties of simple type theory*, Journal of Symbolic Logic 25 (1960), no. 4, pp. 305–326.

[12] Leslie G. Valiant, *Short monotone formulae for the majority function*, Journal of Algorithms 5 (1984), no. 3, pp. 363–366.

[13] Ingo Wegener, *The complexity of Boolean functions*, Teubner, Stuttgart, 1987.