Citation for published version:
Bruscoli, P & Strassburger, L 2017, On the Length of Medial-Switch-Mix Derivations. in J Kennedy & RJGB de Queiroz (eds), Logic, Language and Computation: 24th International Workshop, WoLLIC 2017, London, UK, July 18–21, 2017: Proceedings. Lecture Notes in Computer Science, vol. 10388, Springer Verlag, Berlin, Germany, pp. 68-79. https://doi.org/10.1007/978-3-662-55386-2_5

DOI:
10.1007/978-3-662-55386-2_5

Publication date:
2017

Document Version
Peer reviewed version

Link to publication

The final publication is available at Springer via https://doi.org/10.1007/978-3-662-55386-2_5

On the Length of Medial-Switch-Mix Derivations

Paola Bruscoli[1] and Lutz Straßburger[2]

[1] University of Bath, UK
[2] Inria, France

Abstract. Switch and medial are two inference rules that play a central role in many deep inference proof systems. In specific proof systems, the mix rule may also be present. In this paper we show that the maximal length of a derivation using only the inference rules for switch, medial, and mix, modulo associativity and commutativity of the two binary connectives involved, is quadratic in the size of the formula at the conclusion of the derivation. This shows, at the same time, the termination of the rewrite system.

1 Introduction

Deep inference is a well established methodology in proof theory; it generalises the more traditional proof theoretical methods, while simultaneously improving our ability in studying proofs from the point of view of normalisation and complexity, addressing therefore the problem of proof identity. For the interested reader, the web site `http://alessio.guglielmi.name/res/cos/` provides a detailed overview of the collective developments in deep inference, spanning almost two decades of activities by several research groups and individuals.

For the purposes of this paper, however, we will limit ourselves to recall the essential feature that distinguishes deep inference from the traditional proof formalisms, when describing inference rules of some logical proof system. Deep inference applies logical inference in contexts, i.e. it allows to manipulate formulas at arbitrary depth in any context. In contrast, in the more traditional formalisms, including sequent calculus or natural deduction, the decomposition of a formula around its main connective strictly determines the shape of the inference rules of the proof system, and ultimately the shape of its proofs.

Therefore, on a technical level, deep inference lends itself to be studied also from the perspective of modern term rewriting, and in an easier way than traditional formalisms. This will not come at the expenses of our ability in observing and studying the fundamental proof theoretical properties, including cut-elimination, in a conservative way.

In this paper, we show some upper bounds on the length of deep inference derivations in a (sub)system for classical logic, by taking advantage of both worlds, proof theory and rewriting.

In the standard deep inference proof system KS for classical logic [3], we can find the following two rules

$$\mathsf{s}\,\frac{F\{([A \vee B] \wedge C)\}}{F\{[A \vee (B \wedge C)]\}} \qquad \mathsf{m}\,\frac{F\{[(A \wedge B) \vee (C \wedge D)]\}}{F\{([A \vee C] \wedge [B \vee D])\}} \tag{1}$$

called *switch* and *medial*, respectively, where $F\{\ \}$ stands for an arbitrary (positive) formula context and A, B, C, and D are formula variables. In the system KS, the two rules switch and medial are applied modulo associativity and commutativity of conjunction and disjunction. The switch rule has been well investigated from the proof-theoretic as well as from the category-theoretic point of view because of its important role in linear logic [2]. The properties of the medial rule have originally been investigated in [20].

Switch and medial in combination have been studied in [21,18] from the perspective of semantics of proofs in deep inference. Moreover, the development of atomic flows in [13], centred on proof normalisation, provides a more abstract view of classical proofs by hiding switch-medial steps. A preliminary account from perspective of proof complexity can be found in [5].

In this paper we look at the two rules of switch and medial as rewriting system. Often these two rules may operate in presence of the mix rule, i.e.

$$\mathsf{mix}\,\frac{F\{[A \vee B]\}}{F\{(A \wedge B)\}} \tag{2}$$

that may be induced by the interplay of the units under specific conditions in the proof system, or just be made explicitly available in the system. For classical logic, for example, it is a consequence of weakening.

Our main result, in terms of complexity, is that the length of a derivation using only switch, medial, and mix is bounded by a quadratic function in the size of the conclusion of the derivation. Clearly, any such bound for a system with mix also holds for the system without mix. We will however also include the presentation of a more immediate cubic bound that has independent interest.

2 Rewriting with Switch and Medial (and Mix)

Formulas are generated from a countable set $\mathscr{A} = \{a, b, c, \ldots\}$ of atoms via the binary connectives $\wedge$ and $\vee$, called *and* and *or*, respectively. To ease readability of large formulas we will use [ ] for parentheses around disjunctions and ( ) for parentheses around conjunctions.

To simplify the presentation, we do not use the units $\top$ (*truth*) and $\bot$ (*falsum*) in this paper. It follows from the work by Das [9] that our results would also hold under the presence of the units.

The *size* of a formula A, denoted by σA, is the number of atom occurrences in A:

$$\begin{aligned} \sigma a &= 1 \\ \sigma[A \vee B] &= \sigma A + \sigma B \\ \sigma(A \wedge B) &= \sigma A + \sigma B \end{aligned}$$

Example: $\sigma([a \vee b] \wedge [[a \vee c] \vee b]) = 5$.

The *tree-or-number* of a formula A, denoted by θA, is the number of occurrences of the symbol $\vee$ in A:

$$\begin{aligned} \theta a &= 0 \\ \theta[A \vee B] &= \theta A + \theta B + 1 \\ \theta(A \wedge B) &= \theta A + \theta B \end{aligned}$$

Example: $\theta([a \vee b] \wedge [[a \vee c] \vee b]) = 3$.

The *relweb-or-number* of a formula A, denoted by $\otimes A$, is the number of $\vee$-edges in the relation web [3] of A:

$$\begin{aligned} \otimes a &= 0 \\ \otimes[A \vee B] &= \otimes A + \otimes B + \sigma A \cdot \sigma B \\ \otimes(A \wedge B) &= \otimes A + \otimes B \end{aligned}$$

Example: $\otimes([a \vee b] \wedge [[a \vee c] \vee b]) = 4$.

These values are stable under context application, as shown below.

Lemma 2.1 *Let P and Q be formulas with $\sigma P = \sigma Q$, and let $F\{\ \}$ be a formula context.*

1. *If $\theta P = \theta Q$, then $\theta F\{P\} = \theta F\{Q\}$.*
2. *If $\theta P > \theta Q$, then $\theta F\{P\} > \theta F\{Q\}$.*
3. *If $\otimes P = \otimes Q$, then $\otimes F\{P\} = \otimes F\{Q\}$.*
4. *If $\otimes P > \otimes Q$, then $\otimes F\{P\} > \otimes F\{Q\}$.*

Proof By induction on the structure of $F\{\ \}$. □

Consider the following rewrite rules on formulas:

$$\mathsf{m}\,\frac{[(A \wedge B) \vee (C \wedge D)]}{([A \vee C] \wedge [B \vee D])} \qquad \mathsf{s}\,\frac{([A \vee B] \wedge C)}{[A \vee (B \wedge C)]} \qquad \mathsf{mix}\,\frac{(A \wedge B)}{[A \vee B]} \tag{3}$$

The rules in (3) are written in the style of inference rules in proof theory (premiss on top implies the conclusion at the bottom), but they behave as rewrite rules in term rewriting, i.e., they can be applied inside any formula context $F\{\ \}$.

[3] The relation web of a formula provides a graph-based representation of a formula in deep inference contextual rewriting, as in [12]. An equivalent definition of the relweb-or-number is the number of edges in the cograph for $\vee$.

The rewriting rules in (3) are applied modulo associativity and commutativity for $\wedge$ and $\vee$. More precisely, we will do rewriting modulo the equational theory generated by

$$\begin{array}{ll} (A \wedge (B \wedge C)) = ((A \wedge B) \wedge C) & (A \wedge B) = (B \wedge A) \\ [A \vee [B \vee C]] = [[A \vee B] \vee C] & [A \vee B] = [B \vee A] \end{array} \tag{4}$$

Lemma 2.2 *If $P = Q$, then $\theta P = \theta Q$ and $\circledast P = \circledast Q$.*

Proof Consider the equation of associativity of disjunction. We have

$$\begin{aligned} \theta[A \vee [B \vee C]] &= \theta A + \theta B + \theta C + 2 \\ &= \theta[[A \vee B] \vee C] \end{aligned}$$

and

$$\begin{aligned} \circledast[A \vee [B \vee C]] &= \circledast A + (\circledast B + \circledast C + \sigma B \cdot \sigma C) + \sigma A \cdot (\sigma B + \sigma C) \\ &= \circledast A + \circledast B + \circledast C + \sigma A \cdot \sigma B + \sigma A \cdot \sigma C + \sigma B \cdot \sigma C \\ &= \circledast[[A \vee B] \vee C] \end{aligned}$$

Now apply Lemma 2.1. The other cases are similar (and simpler). □

Lemma 2.3 *Let the rule* $\rho \dfrac{Q}{P}$ *be given.*

1. *If ρ is* m*, then $\theta Q < \theta P$.*
2. *If ρ is* s*, then $\theta Q = \theta P$ and $\circledast Q < \circledast P$.*
3. *If ρ is* mix*, then $\theta Q < \theta P$ and $\circledast Q < \circledast P$.*

Proof 1. Case of medial:

$$\begin{aligned} \theta[(A \wedge B) \vee (C \wedge D)] &= \theta A + \theta B + 1 + \theta C + \theta D + 1 \\ &< \theta A + \theta B + \theta C + \theta D + 1 \\ &= \theta([A \vee C] \wedge [B \vee D]) \end{aligned}$$

2. Case of switch:

$$\begin{aligned} \theta([A \vee B] \wedge C) &= \theta A + \theta B + 1 + \theta C \\ &= \theta[A \vee (B \wedge C)] \end{aligned}$$

and

$$\begin{aligned} \circledast([A \vee B] \wedge C) &= \circledast A + \circledast B + \sigma A \cdot \sigma B + \circledast C \\ &< \circledast A + \circledast B + \circledast C + \sigma A \cdot \sigma B + \sigma A \cdot \sigma C \\ &= \circledast[A \vee (B \wedge C)] \end{aligned}$$

3. Case of mix:

$$\begin{aligned} \theta(A \wedge B) &= \theta A + \theta B \\ &< \theta A + \theta B + 1 \\ &= \theta[A \vee B] \end{aligned} \qquad \text{and} \qquad \begin{aligned} \circledast(A \wedge B) &= \circledast A + \circledast B \\ &< \circledast A + \circledast B + \sigma A \cdot \sigma B \\ &= \circledast[A \vee B] \end{aligned}$$

Now apply Lemma 2.1. □

3 The Cubic Bound

Before showing the quadratic bound, we present a cubic bound that we consider of independent interest for the following reasons. First, the proof of the cubic bound is rather simple and flexible, so it might be of interest for other logics (not just classical logic) especially in relation to aspects of system implementations [17]. Second, the cubic bound on medial-switch-mix derivations has been generalised to arbitrary (sound) linear systems for classical logic [10,11], in the sense that any derivation of super-cubic length must derive a (semantically) trivial inference.

Separating out the different proofs for the cubic and quadratic bound might help to answer the open question whether the quadratic bound can also be generalized, or whether it is truly specific to medial-switch-mix derivations.

For a formula P, define its *MSM-measure*, denoted by $\mathsf{msm}P$, as the pair

$$\mathsf{msm}P = \langle \theta P, \circledast P \rangle \quad .$$

We use the lexicographic ordering for this measure on formulae:

$$\mathsf{msm}Q < \mathsf{msm}P \qquad \text{iff} \qquad \theta Q < \theta P \quad \text{or} \quad (\ \theta Q = \theta P \quad \text{and} \quad \circledast Q < \circledast P\ ) \quad .$$

In the sequel, we assume the common notions and terminology on derivations and proof construction, given a (finite) set of inference rules.

The notation $\begin{array}{c} Q \\ \mathscr{S} \Big\| \varDelta \\ P \end{array}$ stands for a derivation $\varDelta$, with premiss Q and conclusion P, obtained with the rules in $\mathscr{S}$; by $\mathsf{length}(\varDelta)$ we intend the number of instances of rules of $\mathscr{S}$ that have been applied in $\varDelta$.

Proposition 3.1 *Let $\varDelta$ be the following derivation* $\begin{array}{c} Q \\ \{\mathsf{m},\mathsf{s},\mathsf{mix}\} \Big\| \varDelta \\ P \end{array}$, *where $\sigma P = n$. Then,* $\mathsf{length}(\varDelta) < n^3$.

Proof By Lemma 2.2 we have that msm is stable under the equivalence of formulas, and from Lemma 2.3 we can conclude that msm strictly decreases at each step when going bottom-up in the derivation. We also have that $\theta P < n$ and $\circledast P \le n^2$. Hence, $\mathsf{length}(\varDelta) < n \cdot n^2$. □

4 The Quadratic Bound

The analysis that delivers a quadratic bound is performed by treating separately, in a given derivation, those sub-derivations that use only switch rule from those that use both mix and medial. We will keep the presentation slightly more informal to help the intuition.

Lemma 4.1 *Let Δ be the following derivation* $\{\mathsf{m},\mathsf{mix}\}\,\begin{array}{c}Q\\ \Vert\,\Delta\\ P\end{array}$, *where $\sigma P = n$. Then* $\mathsf{length}(\Delta) < n$.

Proof As observed in the proof of Proposition 3.1, we have that $\theta P < n$, and this value strictly decreases with each application of medial and mix. □

Moreover, let γP be the number of $\wedge$-occurrences in a formula P; then, it is trivial to show that

$$\sigma P = \theta P + \gamma P + 1 \quad .$$

In the following, we identify a formula P with its formula-tree, and every node of that tree is identified with the subformula occurrence rooted at that node.

Let $P = S\{R\}$, where R is an $\wedge$-node in P, i.e., $R = (A \wedge B)$ for some A and B. We define the following notions:

- the *switch-potential of R in P* is the number of $\vee$-nodes in the context $S\{\ \}$;
- the *switch-potential of P*, denoted by $\mathsf{sp}P$ is the sum of the switch-potentials of all $\wedge$-nodes in P.

Example:

The formula $([a \vee b] \wedge [[a \vee c] \vee b])$ contains only one $\wedge$-node, and its switch-potential is 0. But the formula $A = [(a \wedge b) \vee ((a \wedge c) \wedge b)]$ has 3 $\wedge$-nodes, each of which has switch-potential 1. Hence $\mathsf{sp}A = 3$.
The switch-potential of formulae is preserved through associativity and commutativity of the two operators and under context closure.

Lemma 4.2 *If P and Q are formulas and $P = Q$ then $\mathsf{sp}P = \mathsf{sp}Q$.*

Proof First note, that whenever $P = Q$ and $\mathsf{sp}P = \mathsf{sp}Q$, then for all contexts $S\{\ \}$ we have $\mathsf{sp}S\{P\} = \mathsf{sp}S\{Q\}$. This can be shown by a straightforward induction on $S\{\ \}$. Hence, it suffices to show that for each equation in (4), the switch-potential of the left-hand side is equal to the switch-potential of the right-hand side. This is straightforward. □

We then consider the switch-potential of premiss and conclusion of the rules switch, medial and mix.

Lemma 4.3 *If* $\mathsf{s}\,\dfrac{Q}{P}$ *is a correct application of the switch rule, then* $\mathsf{sp}Q < \mathsf{sp}P$.

Proof We have that $P = S\{A \vee (B \wedge C)\}$ and $Q = S\{[A \vee B] \wedge C\}$ for some context $S\{\ \}$ and some formulas A, B and C. The $\wedge$-nodes in $S\{\ \}$, in A, and in B do not change their switch-potentials. The switch-potentials of the $\wedge$-nodes in C (if present) are reduced by 1. However, note that the $\wedge$-node $([A \vee B] \wedge C)$ in Q has strictly smaller switch-potential than the $\wedge$-node $(B \wedge C)$ in P. Hence $\mathsf{sp}Q < \mathsf{sp}P$. □

Corollary 4.4 *Let* $\{\mathsf{s}\} \Big\Vert^{Q}_{P} \Delta$ *be given. Then* $\mathsf{length}(\Delta) \leq \gamma P \cdot \theta P < \frac{1}{4}n^2$.

Proof By Lemma 4.3, we immediately get $\mathsf{length}(\Delta) \leq \mathsf{sp}P$. By definition we have $\mathsf{sp}P \leq \gamma P \cdot \theta P$. Since $\gamma P + \theta P = n - 1$, we have $\frac{1}{4}n^2$ as upper bound. □

Observation 4.5 Note that in the derivation Δ in Corollary 4.4 we have $\theta Q = \theta P$ and $\gamma Q = \gamma P$.

Lemma 4.6 *If* $\mathsf{mix}\,\frac{Q}{P}$ *is a correct application of the mix rule, then* $\mathsf{sp}Q < \mathsf{sp}P + \theta P$.

Proof We have that $P = S\{A \vee B\}$ and $Q = S\{A \wedge B\}$ for some context $S\{\ \}$ and some formulas A and B. The switch-potentials of the $\wedge$-nodes in $S\{\ \}$, in A, and in B either remains unchanged by the inference step or is smaller in Q than in P, because one $\vee$-node is removed. However, Q has one $\wedge$-node more than P. Hence, its switch-potential can be increased by the switch-potential of that $\wedge$-node, which is at most θQ. Hence, $\mathsf{sp}Q \leq \mathsf{sp}P + \theta Q$. Since $\theta P = \theta Q + 1$, we get $\mathsf{sp}Q < \mathsf{sp}P + \theta P$. □

Lemma 4.7 *If* $\mathsf{m}\,\frac{Q}{P}$ *is a correct application of the medial rule, then* $\mathsf{sp}Q \leq \mathsf{sp}P + \theta P$.

Proof We have that $P = S\{[A \vee C] \wedge [B \vee D]\}$ and $Q = S\{(A \wedge B) \vee (C \wedge D)\}$ for some context $S\{\ \}$ and some formulas A, B, C and D. The switch-potentials of the $\wedge$-nodes in $S\{\ \}$, and in A, B, C, D could be smaller in Q than in P, because one $\vee$-node is removed. However, one $\wedge$-node in P is replaced by two $\wedge$-nodes in Q, which have both a bigger switch-potential because of the new $\vee$-node as parent. This can be counted as follows: the first new $\wedge$-node in Q increases its switch-potential by 1, compared to the $\wedge$-node in P, whereas the second $\wedge$-node in Q has a switch-potential of at most θQ. Hence $\mathsf{sp}Q \leq \mathsf{sp}P + 1 + \theta Q = \mathsf{sp}P + \theta P$. □

We can now combine these results.

Proposition 4.8 *Let* Δ *be a given derivation* $\{\mathsf{m},\mathsf{s},\mathsf{mix}\} \Big\Vert^{Q}_{P} \Delta$ *where* $\sigma P = n$. *Then* $\mathsf{length}(\Delta) < \frac{1}{2}(n^2 + n)$.

Proof Without loss of generality, we can assume that Δ has shape

$$
\begin{array}{c}
Q_m \\
\{\mathsf{s}\} \Big\| \, \Delta_m \\
\rho_m \dfrac{P_m}{Q_{m-1}} \\
\{\mathsf{s}\} \Big\| \, \Delta_{m-1} \\
\vdots \\
\{\mathsf{s}\} \Big\| \, \Delta_2 \\
\rho_2 \dfrac{P_2}{Q_1} \\
\{\mathsf{s}\} \Big\| \, \Delta_1 \\
\rho_1 \dfrac{P_1}{Q_0} \\
\{\mathsf{s}\} \Big\| \, \Delta_0 \\
P_0
\end{array}
$$

where $P = P_0$, $Q = Q_m$, and where ρ_i is an instance of m or mix, for every $i \in \{1, \ldots, m\}$. Let us consider, for every $i \in \{0, \ldots, m\}$, the numbers of disjunctions and conjunctions in P_i, respectively denoted by d_i and c_i, as follows:

$$d_i = \theta P_i = \theta Q_i \qquad \text{and} \qquad c_i = \gamma P_i = \gamma Q_i \quad .$$

Observe that $d_i + c_i = n-1$. By the same argument as in the proof of Lemma 4.1, we have that $m = c_m - c_0 = d_0 - d_m$. In particular, we have that

$$m < d_0 < n \quad , \tag{5}$$

and for all $i \in \{0, \ldots, m-1\}$ we have that

$$d_{i+1} = d_i - 1 \qquad \text{and} \qquad c_{i+1} = c_i + 1 \quad . \tag{6}$$

We define, for every $i \in \{0, \ldots, m\}$, the switch-potentials as

$$p_i = \mathsf{sp} P_i \qquad \text{and} \qquad q_i = \mathsf{sp} Q_i \quad ,$$

and we have that $p_i \leq d_i c_i$. If we let $l_i = \mathsf{length}(\Delta_i)$, then we obtain

$$l_i \leq p_i - q_i \quad . \tag{7}$$

By Lemmas 4.6 and 4.7 we also have

$$p_{i+1} \leq q_i + d_i \quad . \tag{8}$$

The remainder of the proof is a simple calculation:

$$
\begin{aligned}
\mathsf{length}(\Delta) &= m + l_0 + l_1 + \cdots + l_{m-1} + l_m \\
&\leq m + (p_0 - q_0) + (p_1 - q_1) + \cdots + (p_{m-1} - q_{m-1}) + (p_m - q_m) \\
&\leq m + (d_0 c_0 - q_0) + (q_0 + d_0 - q_1) + \cdots \\
&\qquad\qquad + (q_{m-2} + d_{m-2} - q_{m-1}) + (q_{m-1} + d_{m-1} - q_m) \\
&\leq m + d_0 c_0 + d_0 + d_1 + \cdots + d_{m-2} + d_{m-1} \\
&= m + d_0(n - 1 - d_0) + d_0 + (d_0 - 1) + \cdots \\
&\qquad\qquad + (d_0 - (m-2)) + (d_0 - (m-1)) \\
&= m + d_0(n - 1 - d_0) + m d_0 - \sum_{i=1}^{m-1} i \\
&= m + d_0 n - d_0 - d_0^2 + m d_0 - \frac{m(m-1)}{2} \\
&= \frac{1}{2}(2nd_0 + 2md_0 - 2d_0^2 - m^2 + 3m - 2d_0) \\
&< \frac{1}{2}(n^2 + n) \quad ,
\end{aligned}
$$

where the last inequality follows by observing that

$$
\begin{aligned}
0 &< (n - m) + (d_0 - m) + (d_0 - m) \\
&= n + 2d_0 - 3m \\
&= n - (3m - 2d_0)
\end{aligned}
$$

and

$$
\begin{aligned}
0 &< (n - d_0)^2 + (d_0 - m)^2 \\
&= n^2 - 2nd_0 + d_0^2 + d_0^2 - 2d_0 m + m^2 \\
&= n^2 - (2nd_0 + 2d_0 m - 2d_0^2 - m^2) \quad ,
\end{aligned}
$$

completing, thus, the proof. □

Some remarks are in order, to comment on the bounds that we have obtained in this study.

Remark 4.9 The linear bound given by Lemma 4.1 cannot be reduced. For derivations containing mix, this is obvious, but even when only medial is allowed, we can form the following derivation

$$
\begin{array}{c}
[(a_{11} \wedge a_{21} \wedge \cdots \wedge a_{m1}) \vee \cdots \vee (a_{1k} \wedge a_{2k} \wedge \cdots \wedge a_{mk})] \\
\{\mathsf{m}\} \| \, \Delta \\
([a_{11} \vee a_{12} \vee \cdots \vee a_{1k}] \wedge \cdots \wedge [a_{m1} \vee a_{m2} \vee \cdots \vee a_{mk}])
\end{array}
\tag{9}
$$

that contains $(k-1)(m-1)$ instances of medial. The size of the formulas in Δ is $n = km$. Hence $\mathsf{length}(\Delta) = km - k - m + 1$. If $k = m$, then $\mathsf{length}(\Delta) = n - 2\sqrt{n} + 1$. Lamarche proposes in [18] a matrix notation for denoting the derivation Δ in (9).

Remark 4.10 For derivations that use only the switch rule, the following example shows that the quadratic bound of Lemma 4.4 cannot be pushed further down, and also that the constant factor of $\frac{1}{4}$ is already optimal. Consider the following derivation where only switch is applied:

$$
\begin{array}{c}
((\cdots(([a_1 \vee [a_2 \vee [a_3 \vee \cdots \vee [a_m \vee b]\cdots]]] \wedge c_1) \wedge c_2) \cdots \wedge c_{k-1}) \wedge c_k) \\
\{\mathsf{s}\} \Big\| \ \Delta_1 \\
((\cdots([a_1 \vee [a_2 \vee [a_3 \vee \cdots \vee [a_m \vee (b \wedge c_1)]\cdots]]] \wedge c_2) \cdots \wedge c_{k-1}) \wedge c_k) \\
\{\mathsf{s}\} \Big\| \ \Delta_2 \\
\vdots \\
\{\mathsf{s}\} \Big\| \ \Delta_{k-1} \\
([a_1 \vee [a_2 \vee [a_3 \vee \cdots \vee [a_m \vee (\cdots((b \wedge c_1) \wedge c_2) \cdots \wedge c_{k-1})]\cdots]]] \wedge c_k) \\
\{\mathsf{s}\} \Big\| \ \Delta_k \\
[a_1 \vee [a_2 \vee [a_3 \vee \cdots \vee [a_m \vee ((\cdots((b \wedge c_1) \wedge c_2) \cdots \wedge c_{k-1}) \wedge c_k)]\cdots]]]
\end{array}
\tag{10}
$$

For every $i \in \{1, \ldots, k\}$, we have that each Δ_i consists of m switches. Hence $\mathsf{length}(\Delta) = mk$. If we let $m = k$, we have $n = 2k + 1$ and $\mathsf{length}(\Delta) = k^2 = \frac{1}{4}n^2 - \frac{1}{2}n + \frac{1}{4}$.

Remark 4.11 The previous remark also shows that the quadratic bound of Proposition 4.8 cannot be improved. However, it is not known whether the constant factor $\frac{1}{2}$ can be improved (although we know that it must be $\geq \frac{1}{4}$.)

5 Conclusions

Earlier versions of this paper exist since 2008, for research primarily motivated by the need of better understanding the role and shape of the medial rule, especially from the perspective of normalisation. The medial rule is in fact needed in deep inference systems for classical logic to obtain an atomic contraction rule, which, in turns, contributes a form of atomic sharing that influences both normalisation and complexity.

Over time, we noted that variants of the medial rule appear, possibly in disguise, in several different logics of the linear kind, including those with non-commutative operators. It is therefore appropriate studying switch-medial-mix derivations independently from the specific units of the logic and their associated equations. In this sense, the approach based on term rewriting adds an element of generality that could result useful also for aspects of implementations.

The switch-medial-mix fragment is at the core of several investigations from different perspectives, including also the development of atomic flows [13,14] and of the atomic lambda calculus [16]. In particular, atomic flows provide an abstract view of classical derivations by making switch and medial unobservable (hence indirectly related to this topic of study) and enabling the discovery of interesting transformations from the complexity perspective, such as [6], [7] and [8].

The results of this paper have supported well also the study of the length of derivations consisting only of linear inferences, as developed in [9], and more recently in [10] and [11]. The two last works show that our cubic bound also holds for non-trivial derivations made of arbitrary sound linear inferences, not just switch and medial. It still remains an open question whether the general case also has a quadratic bound.

As a remark, our medial-switch-mix system differs from the system studied in [1], which contains a rule that is the "inverse" of our medial rule, with premiss and conclusion swapped around.

As a matter of fact, an interesting line of enquiry is what happens when we combine various forms of switch and medial for various connectives. Examples of such a combination is the local system for linear logic [19], and the systems that extend the basic BV that contains a sequential operator inspired by process algebras [15,4]. More recently, a very exciting development that generalises switch and medial through the use of *subatomic logic* is in [22].

For the central role that switch and medial rules have in all deep inference systems, and for the richness of results collected from different perspectives that confirm also our bounds, we hope that this paper proves useful also in relation to implementations as well as further studies in complexity.

6 Acknowledgements

We are grateful to Alessio Guglielmi and Tom Gundersen, for their encouragements and checks, and to the anonymous referees. Paola Bruscoli was funded by EPSRC Project EP/K018868/1 "Efficient and Natural Proof Systems. This research is also supported by ANR Project FISP – "The Fine Structure of Formal Proof Systems and their Computational Interpretations."

References

1. Denis Bechet, Philippe de Groote, and Christian Retoré. A complete axiomatisation of the inclusion of series-parallel partial orders. In H. Common, editor, *Rewriting Techniques and Applications, RTA 1997*, volume 1232 of *LNCS*, pages 230–240. Springer, 1997.
2. Richard Blute, Robin Cockett, Robert Seely, and Todd Trimble. Natural deduction and coherence for weakly distributive categories. *J. of Pure and Applied Algebra*, 113:229–296, 1996.
3. Kai Brünnler and Alwen Fernanto Tiu. A local system for classical logic. In R. Nieuwenhuis and A. Voronkov, editors, *LPAR 2001*, volume 2250 of *LNAI*, pages 347–361. Springer, 2001.
4. 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 Artificial Intelligence*, pages 302–316. Springer-Verlag, 2002.
5. Paola Bruscoli and Alessio Guglielmi. On the proof complexity of deep inference. *ACM Transactions on Computational Logic*, 10(2):1–34, 2009. Article 14.

6. Paola Bruscoli, Alessio Guglielmi, Tom Gundersen, and Michel Parigot. A quasipolynomial cut-elimination procedure in deep inference via atomic flows and threshold formulae. In *LPAR-16*, volume 6355 of *LNCS*, pages 136–153. Springer-Verlag, 2010.
7. Paola Bruscoli, Alessio Guglielmi, Tom Gundersen, and Michel Parigot. A quasipolynomial normalisation in deep inference via atomic flows and threshold formulae. *Logical Methods in Computer Science*, 12((1:5)):1–30, 2016.
8. Anupam Das. Complexity of deep inference via atomic flows. In S. Barry Cooper, Anuj Dawar, and Benedikt Löwe, editors, *Computability in Europe*, volume 7318 of *LNCS*, pages 139–150. Springer-Verlag, 2012.
9. Anupam Das. Rewriting with linear inferences in propositional logic. In Femke van Raamsdonk, editor, *24th International Conference on Rewriting Techniques and Applications (RTA)*, volume 21 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 158–173. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2013.
10. Anupam Das and Lutz Straßburger. No complete linear term rewriting system for propositional logic. In Maribel Fernández, editor, *26th International Conference on Rewriting Techniques and Applications, RTA 2015, June 29 to July 1, 2015, Warsaw, Poland*, volume 36 of *LIPIcs*, pages 127–142, 2015.
11. Anupam Das and Lutz Straßburger. On linear rewriting systems for boolean logic and some applications to proof theory. *Logical Methods in Computer Science*, 12(4), 2016.
12. Alessio Guglielmi. A system of interaction and structure. *ACM Transactions on Computational Logic*, 8(1):1–64, 2007.
13. Alessio Guglielmi and Tom Gundersen. Normalisation control in deep inference via atomic flows. *Logical Methods in Computer Science*, 4(1:9):1–36, 2008.
14. Alessio Guglielmi, Tom Gundersen, and Lutz Straßburger. Breaking paths in atomic flows for classical logic. In *LICS 2010*, 2010.
15. Alessio Guglielmi and Lutz Straßburger. Non-commutativity and MELL in the calculus of structures. In Laurent Fribourg, editor, *Computer Science Logic, CSL 2001*, volume 2142 of *LNCS*, pages 54–68. Springer-Verlag, 2001.
16. Tom Gundersen, Willem Heijltjes, and Michel Parigot. Atomic lambda calculus: A typed lambda-calculus with explicit sharing. In *28th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2013, New Orleans, LA, USA, June 25-28, 2013*, pages 311–320. IEEE Computer Society, 2013.
17. Ozan Kahramanogullari. Interaction and depth against nondeterminism in proof search. *Logical Methods in Computer Science*, 10(2), 2014.
18. François Lamarche. Exploring the gap between linear and classical logic. *Theory and Applications of Categories*, 18(18):473–535, 2007.
19. Lutz Straßburger. A local system for linear logic. In Matthias Baaz and Andrei Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning, LPAR 2002*, volume 2514 of *LNAI*, pages 388–402. Springer-Verlag, 2002.
20. Lutz Straßburger. A characterisation of medial as rewriting rule. In Franz Baader, editor, *Term Rewriting and Applications, RTA'07*, volume 4533 of *LNCS*, pages 344–358. Springer, 2007.
21. Lutz Straßburger. On the axiomatisation of Boolean categories with and without medial. *Theory and Applications of Categories*, 18(18):536–601, 2007.
22. Andrea Aler Tubella. *A study of normalisation through subatomic logic*. PhD thesis, University of Bath, 2016.