A Nonstandard Standardization Theorem

Beniamino Accattoli
Carnegie Mellon University
United States

Eduardo Bonelli
Univ. Nac. de Quilmes and CONICET
Argentina

Delia Kesner
Univ. Paris-Diderot, SPC, PPS, CNRS
France

Carlos Lombardi
Univ. Nac. de Quilmes and Univ. de Buenos Aires
Univ. Paris-Diderot, SPC, PPS, CNRS
Argentina and France

Abstract

Standardization is a fundamental notion for connecting programming languages and rewriting calculi. Since both programming languages and calculi rely on substitution for defining their dynamics, explicit substitutions (ES) help further close the gap between theory and practice.

This paper focuses on standardization for the linear substitution calculus, a calculus with ES capable of mimicking reduction in \( \lambda \)-calculus and linear logic proof-nets. For the latter, proof-nets can be formalized by means of a simple equational theory over the linear substitution calculus.

Contrary to other extant calculi with ES, our system can be equipped with a residual theory in the sense of Lévy, which is used to prove a left-to-right standardization theorem for the calculus with ES but without the equational theory. Such a theorem, however, does not lift from the calculus with ES to proof-nets, because the notion of left-to-right derivation is not preserved by the equational theory. We then relax the notion of left-to-right standard derivation, based on a total order on redexes, to a more liberal notion of standard derivation based on partial orders.

Our proofs rely on Gonthier, Lévy, and Melliès’ axiomatic theory for standardization. However, we go beyond merely applying their framework, revisiting some of its key concepts: we obtain uniqueness (modulo) of standard derivations in an abstract way and we provide a coinductive characterization of their key abstract notion of external redex. This last point is then used to give a simple proof that linear head reduction—a nondeterministic strategy having a central role in the theory of linear logic—is standard.

Categories and Subject descriptors F.3.2 [Logics and Meanings of Programs]: Semantics of Programming Languages - Operational semantics; F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic - Lambda calculus and related systems; I.1.3 [Symbolic and Algebraic Manipulation]: Languages and Systems - evaluation strategies

General Terms Theory, Languages.

Keywords Lambda-calculus, Standardization, Residuals, Explicit Substitutions, Linear Logic, Proof-Nets.

1. Introduction

In his seminal paper [35] Plotkin introduced the idea of relating a calculus (given by means of an equational theory induced by a rewriting relation) and a programming language (specified as an abstract machine) via a standardization theorem: the programming language implements the standard strategy associated to the calculus. Let us recall what standardization is about. The idea is to identify a class of derivations, called standard, that is complete: whenever \( t \rightarrow u \) then there is a standard derivation from \( t \) to \( u \). Often, it is possible to describe a reduction strategy that produces standard derivations only (the leftmost-outmost strategy in the \( \lambda \)-calculus), and for which the normalization property holds as a corollary of the standardization theorem: if \( t \) has a normal form then the standard strategy will reach this normal form. Completeness and normalization are the justifications for Plotkin’s approach.

Another approach to close the gap between calculi and programming languages is to resort to explicit substitutions (ES), that are refinements of \( \lambda \)-calculus where evaluation is decomposed into small steps, and that can be thought as a framework in which to specify abstract machines as strategies.

It would be natural to expect calculi with ES to enjoy standardization theorems whose induced standard strategies justify abstract machines. Surprisingly, the literature does not present any such relationship. This is due to some inherent difficulties. Calculi with ES are complex and fragile rewriting systems for which already studying normalization is technically demanding. Standardization—that mixes confluence and termination arguments—generally is simply out of scope. In particular, ES calculi present critical pairs and thus lack orthogonality, a basic syntactic property enjoyed by the \( \lambda \)-calculus and on which most standardization techniques rely on.

This paper makes significant progress on standardization for ES. The starting point is a recent advance in the field of ES. A new generation of calculi with ES at a distance, having very simple meta-theories and conceived as behaviorally equivalent copies of graphical languages such as linear logic proof-nets, have recently been introduced by Accattoli and Kesner in [6] (see also [3–5]).
Traditional calculi with ES implement the ordinary substitution by percolating ES through the term structure until they reach variable occurrences, on which they finally substitute or get garbage collected. The key feature of distance calculi is that the percolating process is factored out from the computational process. Indeed, ES at a distance do not move: their dynamics is defined using contexts (i.e., terms with holes) that allows them to act directly on single variable occurrences (via a context isolating the occurrence), without any need to commute with the other constructors in between.

The Framework. In this paper we focus on a specific calculus at a distance, the linear substitution calculus $\lambda_{sub}$, that is both a slight generalization of a calculus by Robin Milner [33] (related to bigraphs), from which it inherits the slight generalization of a calculus by Robin Milner [33] (related to bigraphs), from which it inherits the substitution rules at a distance, and a slight modification of the structural $\lambda$-calculus by Accattoli and Kesner [6] (related to proof nets). Such a calculus has some relevant features:

1. **Linear logic and concurrency:** it is tightly connected with the translations of $\lambda$-calculus into linear logic [2] and the $\pi$-calculus [4]. In particular, reduction rules for ES act as exponential cut-elimination rules in linear logic (resp. replication in the $\pi$-calculus).

2. **Linear head reduction (LHR):** given the deep connection between $\lambda_{sub}$ and proof-nets [2], our calculus naturally expresses linear head reduction, a notion of evaluation for proof nets that is connected with other fundamental models [4, 5, 14, 15, 17, 29]. Notably, LHR is a strategy for $\lambda_{sub}$, while it cannot be expressed as a strategy in $\lambda$-calculus.

3. **Simplicity and expressiveness:** the calculus is simple (only 3 rewriting rules) and it enjoys numerous desirable properties including: a tight correspondence with the lambda calculus (simulating $\beta$-reduction), confluence on terms and metaterms, preservation of strong $\beta$-normalisation and—as we will show— a theory of residuals.

The linear substitution calculus—as do all other extant calculi with ES—presents critical pairs, and so standardization is non-trivial. However, $\lambda_{sub}$ enjoys a unique property, that we call semantical orthogonality.

**Orthogonality.** Orthogonal (first and higher-order) term rewriting systems are defined as left-linear systems without critical pairs. They are the most studied class of rewriting systems, including in particular the $\lambda$-calculus. This notion of syntactic orthogonality is handy and simple, but it has the drawback of being too restrictive. Fortunately, orthogonality can be defined in a more abstract way, as the fact that residuals—a standard concept in rewriting theory (see Section 3)—are sufficient to close local confluence diagrams. This form of semantical orthogonality is a particular property of syntactic orthogonality, but it is of a behavioral nature and more general. The crucial point is that $\lambda_{sub}$ is not syntactically orthogonal but—in contrast to any other calculus with ES—is semantically orthogonal, as we prove in this paper. In this sense $\lambda_{sub}$ is the first calculus with ES that conservatively refines $\lambda$-calculus, preserving its orthogonality. Unfortunately, most techniques for standardization rely on syntactical orthogonality. An exception is the abstract theory of standardization developed by Gonthier, Lévy, and Melliès (GLM) [19], and later refined by Melliès [30, 32]. Such a theory can cope even with non semantically orthogonal calculi as the $\sigma$-calculus, but it takes a much simpler form in presence of semantical orthogonality.

**Left-to-Right Standardization and Beyond.** Two first contributions of this paper are 1) a theory of residuals for $\lambda_{sub}$ and 2) a proof of the axioms of GLM’s theory. These technical results allow us to obtain a left-to-right standardization theorem generalizing the well-known standardization theorem for $\lambda$-calculus, and capturing the deterministic weak (i.e., not reducing under abstractions) variant of LHR—that is the variant that actually matches evaluation in the Krivine Abstract Machine and the $\pi$-calculus—as a standard strategy. It is pleasing that the properties of left-to-right derivations extend naturally to $\lambda_{sub}$. However, such a standardization theorem is not fully satisfying.

A first reason is that LHR (without the weak restriction) is a nondeterministic strategy whose sequences are not necessarily left-to-right standard. However, the nondeterminism of LHR does not affect the final result nor the length of evaluation sequences (technically, LHR enjoys the diamond property). It is then disappointing that LHR is not left-to-right standard, because LHR is a crucial notion in the theory of linear logic.

A second reason for not being satisfied with the left-to-right standardization theorem concerns the close relationship between the linear substitution calculus and linear logic proof-nets, the model behind its inception. Terms and proof-nets are behaviorally equivalent in a strong sense: every term $t$ maps to a proof-net $P_t$, and every evaluation step on $t$ or $P_t$ maps to an evaluation step on the other. Additionally, the redexes are in bijection and so concepts such as residuals transfer from terms to proof-nets and viceversa. The interest of proof-nets is that they provide a quotient of terms; remarkably, this quotient can be explicitly characterized by a simple equational theory $\sim$ (in the style of [6, 23], and here generated by 3 equations only) turning the behavioral equivalence with linear logic proof-nets into a true isomorphism: for every proof-net $P$ there is a $\sim$-class containing all and only the terms behaviorally equivalent to $P$. Consequently, any notion or result that is stable by equivalence $\sim$ immediately lifts to proof-nets. Unfortunately, left-to-right derivations in the linear substitution calculus are not stable by $\sim$, because $\sim$ swaps some constructors, inverting the relative position of some redexes. Thus, our left-to-right standardization theorem does not hold for proof-nets. This is quite disappointing because proof nets provide an operational model, and one would like to have a notion of standard derivation that can be freely transported from the language to the model and viceversa. We then refine our left-to-right standardization theorem.

**Partial Standardization Orders.** Standard derivations are defined as those reductions that respect a certain order on redexes, for instance the left-to-right order. In GLM’s theory such a total order is replaced by a partial order, providing a more general but also subtler setting. To achieve a notion of standard derivation that is stable by $\sim$, we are forced to relax the total left-to-right order into a partial order. In the setting of total orders, given a derivation $\rho : t \rightarrow u$, there is a unique standard derivation $\rho'$ to which $\rho$ standardizes, and moreover a deterministic and normalizing standard strategy can be easily obtained by selecting the minimum redex according to the order. When the order is partial, instead, there can be many standard derivations to which a derivation standardizes, i.e. uniqueness is lost: a standard derivation is an ordered derivation but only up to swaps of $\preceq$-disjoint steps. Moreover, although GLM’s theory does support partial orders, it requires a number of axioms to hold, some of which are not enjoyed by $\lambda_{sub}$ modulo $\sim$. So we shall switch to the axiomatics of Melliès’ PhD thesis [30], which is more general and allows us to prove existence of standard derivations for a suitable partial order. However, alas, $\lambda_{sub}$ modulo $\sim$ does not even satisfy the axioms of [30] required to obtain uniqueness of standard derivations.

1 We distinguish between the leftmost strategy (that repeatedly reduces the leftmost redex) from left-to-right standardization that re-arranges redexes (not necessarily involving the leftmost redex) from left to right. We also use leftmost (considering terms as strings of symbols) for what is sometimes called leftmost-outermost (that is relative to terms seen as trees).
dard derivations. Nor does it supply us with the guarantee that reducing external redexes—the generalization to partial orders of the notion of minimum redexes in the setting of total orders—leads to normal forms. We circumvent these inadequacies of the axiomatic framework by exploiting the embedding of a partial order into a total order, thus proving uniqueness (modulo swappings) of standard derivations and the normalization of a (specific) external strategy.

**Main Contributions.** A summary of our results follows:

1. **The Standardization Theorem.** We identify a partial order ≼_db that is related to the concept of exponential box in linear logic proof-nets, and that is stable by ~. We show that Mellies’ axioms for existence of a standard derivation hold for ≼_db, obtaining a standardization theorem for λ_{db} modulo ~.

2. **Uniqueness Modulo.** For partial orders, GLM’s theory has some additional axioms that ensures uniqueness modulo independent swaps (see Section 8) but one of these axioms does not hold for λ_{db}. Nonetheless, we prove uniqueness in an abstract way, relying on the embedding of the partial box order into the total left-to-right order.

3. **Coinductive External Redexes.** In GLM’s theory external redexes are defined somewhat indirectly, via an extraction process [30]. We clarify this concept by providing a direct and simple coinductive characterization: an external redex is a minimal redex that is persistently minimal, i.e. whose (unique) residual after any other redex is still external.

4. **Applications.** We use the coinductive characterization of external redexes to obtain a normalization theorem for the leftmost strategy and to give a concise proof by coinduction of the fact that LHR is a standard strategy.

The nonstandard character of our standardization result comes at least from the following facts: highlight of the first semantical orthogonal calculus with ES being isomorphic to linear logic proof-nets and enjoying standardization; use of a partial (vs total) order on redexes being stable by equivalence classes; proof of uniqueness (modulo swappings) of standard derivations and an original application of GLM’s theory to a non-syntactical orthogonal calculus with ES being isomorphic to linear logic proof-nets, and that is related to the concept of exponential box in linear logic proof-nets.

**Plan of the Paper.** The next section introduces the linear substitution calculus. Section 3 defines residuals and Section 4 shows that residuals are compatible with the equivalence ~ in a very strong sense. Section 5 explains the ideas behind the abstract approach to standardization and Section 6 proves the left-to-right standardization theorem. Section 7 explains how the left-to-right order can be relaxed to cope with the equational theory and Section 8 proves our equational standardization theorem. Section 9 provides uniqueness of standard derivations. Section 10 presents a coinductive characterization of external redexes and Section 11 proves the normalization theorem and the fact that (weak) linear head reduction is standard.

### 2. The Linear Substitution Calculus

The set of terms of the **linear substitution calculus**, denoted by \( T \), is generated by the following grammar:

\[
-t ::= x \mid tt \mid \lambda x.t \mid t[x/t]
\]

A term \( x \) is called a **variable**, \( t[u] \) an **application**, \( \lambda x.t \) an **abstraction** and \( t[x/u] \) an **explicit substitution**. The notions of free and bound variables are defined as usual plus \( \forall v(t[v/u]) := t[t[v/u] \setminus x] \cup \forall v(u) \). We work with the standard notion of \( \alpha \)-conversion (i.e. renaming of bound variables for abstractions and substitutions). We use \( C \) to denote a **context** \( w.r.t. \) the previous grammar (i.e. a term with a unique occurrence of a designated symbol \( \Box \) called the **hole**). We write \( C[t] \) for the term obtained by replacing the hole of \( C \) by the term \( t \). We write \( C[u] \) when the free variables of \( u \) are not captured by the context \( C \), i.e. there are no abstractions or explicit substitutions in \( C \) that bind the variables of \( f(v) \).

The **λ_{db}-calculus** is given by the set of terms \( T \) and by the **reduction relation** \( \rightarrow_{\lambda_{db}} \) as defined by \( \rightarrow_{db} \), \( \rightarrow_{\alpha} \), \( \rightarrow_{\beta} \), and \( \rightarrow_{\gamma_{ec}} \), which are the closure by contexts \( C \) of the following rewriting rules, where \( L \) denotes a (possibly empty) list of substitutions \( \{x_1/t_1\} \ldots \{x_k/t_k\} \):

\[
(\lambda x.t)Lu \rightarrow_{\alpha} t[x/u]L \quad C[x/u] \rightarrow_{\beta} C[u/x] \quad t[x/u] \rightarrow_{\gamma_{ec}} t \quad \text{if } x \notin \forall v(t)
\]

The names \( db, \lambda, \) and \( \gamma_{ec} \) stand for **distant beta**, **linear substitution**, and **garbage collection**, respectively. Rule \( \rightarrow_{\alpha} \) (resp. \( \rightarrow_{\beta} \)) comes from the structural \( \lambda \)-calculus \([6]\) (resp. M"{u}ller’s calculus \([33]\)), while \( \rightarrow_{\gamma_{ec}} \) belongs to both calculi. In \( db \) we may assume w.l.o.g. that \( \bigcup_{i=1}^{k} \{x_i\} \cap \forall v(u) = \emptyset \) and \( x \notin \forall v(u) \cup \bigcup_{i=1}^{k} \forall v(t_i) \).

Note that the meta-notation \( L = [x_1/t_1] \ldots [x_k/t_k] \) can also be seen as a context \( [x_1/t_1] \ldots [x_k/t_k] \). This fact, together with

\( \tilde{\phi} \) Such a standardization theorem is folklore in the linear logic community. See \([3, 16]\) for details.
the use of a context $C$ in the second rule, and the global side condition in the third rule, justify the idea of rewriting rules at a distance. The substitution context in rule $db$ is motivated by its encoding in proof-nets, where explicit substitutions are partially free to float (i.e. to traverse some term constructors). Such freedom is formalized by the forthcoming graphical equivalence and by the study of its properties in Section 4.

The linear substitution calculus enjoys all the properties required of calculi with ES (including simulation of $\beta$-reduction, preservation of strong normalisation, confluence on terms and metaterms and full composition), whose proofs are simple and omitted, as they are minor variations over those for Milner’s calculus [24], or those of the structural $\lambda$-calculus [6].

In order to study residuals we need to fix a precise terminology about redexes. A redex occurrence in a term $t$ is either a tuple $\langle D, r \rangle$ where $t = D[r]$ and $r = (\lambda x.s)[Lu]$ (a $db$-redex), or a tuple $\langle D, r, C \rangle$ where $t = D[r]$ and $r = C[x/u]$ (a $\alpha$-redex), or a tuple $\langle D, r \rangle$ where $t = D[r]$, $r = [x/u]$ and $x \notin \mathcal{F}(s)$ (a gen-redex). For example, the term $t = (x)(y)[y/z]$ has two different redex occurrences, namely $\langle D_1, r_1, C_1 \rangle = (\langle (x)(y)[y/z], (x)(y) \rangle)$ and $\langle D_2, r_2, C_2 \rangle = (\langle (x)(y)[x/y], (x)(y) \rangle)$. The pattern of a redex is $r$, the second component of the tuple; the box $x$ of the redex is the subterm of the pattern noted $x$. We use $A, B, \ldots$ for redexes occurrences and $\mathcal{R}(t)$ for the set of all redex occurrences of $t$.

Given a redex occurrence $A \in \mathcal{R}(s)$ we write $s \xrightarrow{A} t$ for the reduction step obtained by contracting $A$ in $s$.

The graphical equivalence is given by the contextual, transitive, symmetric and reflexive closure of $\alpha$-conversion (i.e. renaming of bound variables) and the following axioms:

\[
\begin{align*}
& (\lambda y.t)[u/v] \equiv_{ctx} \lambda y.\lambda u.t[u/v] \quad x \notin \mathcal{F}(v) \land y \notin \mathcal{F}(u) \\
& (\lambda y.t)[u/v] \equiv_{\alpha}, \lambda y.\lambda u.t[u/v] \quad y \notin \mathcal{F}(u) \\
& ((t)v)[u/v] \equiv_{\alpha}, \lambda u.t[u/v]v \quad x \notin \mathcal{F}(v) \\
& ((t)v)[u/v] \equiv_{\alpha}, \lambda u.t[u/v]v \quad x \notin \mathcal{F}(v)
\end{align*}
\]

This equivalence characterizes exactly the representation of terms as proof-nets, in the sense that $\approx$ is syntactically orthogonal, a property that no other calculus with ES enjoys. In order to give an intuition on such a phenomenon let us consider a calculus with ES such as $\lambda \times$ [1] or $\lambda x$ [9] containing at least the following reduction rules:

\[
\begin{align*}
& (\lambda x.t)[u/v] \rightarrow_b t[x/u] \\
& ((t)[v/u]) \rightarrow_\alpha t[x/u]v[x/v]
\end{align*}
\]

The following critical pair arises:

\[
(\lambda x.t)[y/v][u/y] \rightarrow_b ((\lambda x.t)[u/v]) \rightarrow_b t[x/u][y/v]
\]

However, the $b$-step has no residual after the $b$-step. The $b$-step also has no residual (note how it would have a residual if the rule were at a distance). The diagram can be closed, but only reducing $\lambda \times$-redexes, and so the calculus is not semantically orthogonal.

The formal developments of residuals is based on a notion of (well-labeled term to be introduced next). Note that since $\lambda x\text{sub}$ is not syntactically orthogonal the technique of underlining the redex pattern [26] cannot be applied, e.g. in $(x)(y)[x/a]$: the underlining does not distinguish between the two occurrences of the $\lambda x$ redexes.

**Labels.** In order to follow redexes along a derivation we mark them with special symbols called labels, denoted $\alpha, \beta, \gamma, \ldots$. The obtained set of labeled terms, denoted by $\mathcal{L}_\mathcal{C}$ is generated by the following grammar.

\[
t ::= x \mid x^{\alpha} \mid tt \mid \lambda x.t \mid t[x/t] \mid t[x^0/t]
\]

The notations $x^{(\alpha)}$, $\lambda x^{(\alpha)}$, and $t[x^{(\alpha)}/t]$ mean that $x$ may or may not be labeled. We write $\mathcal{L}(t)$ to denote the set of all the labels of $t$ and $t^{\alpha}$ to denote the term obtained from $t$ by removing all its labels. Thus for example $(x^0[y^0]y^0/\lambda x^0 z^0)$ is $(x)\langle y, z, \lambda \rangle$.

We extend the meta-notation $\mathcal{L}$ to lists of possibly labeled substitutions, and $\mathcal{C}$ to possibly labeled contexts. Similarly, the notions of free and bound variables are extended to labeled terms as expected together with their corresponding notion of $\alpha$-conversion.

We use $\mathcal{F}(t)$ to denote the subset of $\mathcal{F}(t)$ having at least one labeled occurrence, e.g. $\mathcal{F}(x^{(\alpha)}[y/z]) = \{x\}$.

A labeled reduction $\xrightarrow{\alpha}$ on labeled terms is defined as the contextual closure of the following rewriting rules:

\[
\begin{align*}
& (\lambda x^\alpha.t)[u] \xrightarrow{\alpha db} \lambda y.t[x/u] \\
& C[x^\alpha/u][x/u] \xrightarrow{\alpha ls} C[u][x/u] \\
& t[x^\alpha/u] \xrightarrow{\alpha gc} t \quad x \notin \mathcal{F}(t)
\end{align*}
\]

Definition of residuals for terms naturally extends to labeled ones. A labeled redex $A$ is a redex having a pattern of the form $(\lambda x^\alpha.t)[Lu]$, $C[x^\alpha/u][x/u]$ or $t[x^\alpha/u]$, and $\alpha$ is called the label of the redex $A$. The anchor of a redex (labeled or not) is the variable possibly carrying its label. We will usually associate the labels $\alpha, \beta$ and $\gamma$ to the redex names $A, B$ and $C$ respectively; occasionally we will write $A_\lambda$ to emphasize that $\alpha$ is the label of the redex $A$. We write $\mathcal{R}(A_\lambda)$ for all the redexes of $t$ labeled with $\alpha$.

In order to show some key properties required by the axiomatic approach (cfr. Section 6) we will work with a subset of labeled terms, written $\mathcal{F}(t)$, called well-labeled terms, and defined by:

\[
\begin{itemize}
\item $x \in \mathcal{F}(t)$ and $x^\alpha \in \mathcal{F}(t)$
\end{itemize}
\]

\textbf{ORTHOGONAL.} The subtlety in the study of residuals is that redexes may be duplicated or erased along the way and duplication may even nest two residuals of the same redex. For further references the reader may consult [20, 27, 31, 38]. Another simpler notion of orthogonality is syntactical orthogonality [22], and happens when the system is left-linear and has no critical pairs. Syntactical orthogonality implies semantic orthogonality but the converse does not hold. In particular, the $\lambda x\text{sub}$-calculus is not syntactically orthogonal, but, as we will show in Section 6, it turns out to be semantically orthogonal, a property that no other calculus with ES enjoys.
If \( t \in T_{\text{WC}} \) and \( x \notin \text{fv}(t) \) then \( \lambda x.t \in T_{\text{WC}} \)

If \( t, u \in T_{\text{WC}} \), then \( tu \in T_{\text{WC}} \)

If \( (\lambda x.t)Lu \in T_{\text{WC}} \), then \( (\lambda x.a.t)Lu \in T_{\text{WC}} \)

If \( t,u \in T_{\text{WC}} \), then \( t[x/u] \in T_{\text{WC}} \)

If \( t,u \in T_{\text{WC}} \) and \( x \notin \text{fv}(t) \), then \( t[x'/u] \in T_{\text{WC}} \).

Note that \( \lambda x.a, x, \lambda x.a' \) and \( x[a'/u] \) are not in \( T_{\text{WC}} \). Note also that subterms of well-labeled terms are not necessarily well-labeled (e.g. the abstraction of a labeled db-redex). Well-labeled terms are stable by reduction and graphical equivalence:

**Lemma 1.** Let \( t \in T_{\text{WC}} \). If \( t \xrightarrow{\alpha} u \) or \( t \sim u \), then \( u \in T_{\text{WC}} \).

### Residuals

Here we define residuals for the reduction relation \( \lambda_{\text{sub}} \). In Section 4 we will extend them to the \( \lambda_{\text{sub}} \)-calculus.

A term \( t \) can be labeled in different ways, leading to different variants of \( u \). More precisely, we say that \( t \) is a **variant** of \( u \) iff \( t^0 = u^0 \). Thus in particular, \( t \) is a variant of itself. If \( t \) is a variant of \( u \), then we will consider the **obvious bijection** between the sets of redexes of \( t \) and \( u \); sometimes, we will even identify some particular redexes of them which are related by this bijection.

To define the residuals of a redex \( A \in \text{Red}(t) \) after a step \( t \text{ } B \text{ } u \), we consider the \( \alpha \)-lift of \( t \) w.r.t. \( A \) by \( \alpha \), written \( \text{lift}(t, A, \alpha) \), which is a particular variant of \( t \) obtained as follows. Let \( \alpha \) be a fresh symbol, i.e. \( \alpha \notin \text{Lab}(t) \). If the redex \( A \) is already labeled in \( t \), then change the label of \( A \) to \( \alpha \), otherwise, if \( A \) is not already labeled, assign \( \alpha \) to \( A \). Let us write \( \text{lift}(t, A, \alpha) \) for the induced associated reduction step. Then, the **set of residuals of \( A \)** after \( B \) is given by \( \text{Red}(B) := \{ \text{Red}_d(u) \mid \text{lift}(t, A, \alpha) \} \)

It is clear that this definition is completely **independent** from the variant used to lift the term \( t \). We write \( \text{Red}(t) \) for \( A \in \text{Red}(t) \), and \( \text{Red}(B) \) for \( \bigcup_{A \in \text{Red}(t)} A \).

For example, taking \( v = (x^0 x^0 y^0)(x/y) \), \( B = (\square, u, x^0 \square x^0) \) (so that \( v \xrightarrow{\lambda} (x^0 y^0 x^0 x/y) = v' \) and \( A = \text{Red}_d(v) \setminus B \), we have \( \text{Red}(B) = \{ B', B'' \} \), where \( B' = (\square, v', x^0 y^0) \) and \( B'' = (\square, v, x^0 y^0) \).

**Creation.** Given a reduction step \( t \xrightarrow{\alpha} A \), the set of redexes in \( u \) that are not residuals of \( t \) in \( u \) are said to be **created**. While there are only 3 possible ways to create redexes in \( \lambda \)-calculus [27], in the \( \lambda_{\text{sub}} \)-calculus redexes may be created in one of 6 possible ways, we just show below one example for each case, where the reduced redex has label \( \alpha \) and the created redex is underlined:

- **db creates** \( \langle (\lambda x.a, x)(y,s) \rangle \) \( v \xrightarrow{\text{db}} \langle (y,s)[x/u] \rangle \)
- **db creates 1s** \( \langle \lambda x.a \rangle \) \( v \xrightarrow{\text{db}} x[x/u] \)
- **db creates gc** \( \langle \lambda x.a, y \rangle \) \( v \xrightarrow{\text{db}} y[y/u] \)
- **1s creates db** \( x^0 \) \( v \xrightarrow{\text{1s}} \langle (y,s)[x/\lambda y.s] \rangle \)
- **1s creates gc** \( x^0 \) \( v \xrightarrow{\text{1s}} y[y/u] \)
- **gc creates gc** \( z[y^0] \) \( v \xrightarrow{\text{gc}} z[x/u] \)

Note that each step creates a redex of each kind, 1s-steps can create \( \{ \text{db, gc} \} \)-redexes, and gc-redexes can only create other gc-redexes. Note also that whenever a labeled redex creates a redex \( A \), then \( A \) is not labeled.

**Developments.** Labels are particularly useful to study **developments**, i.e. derivations in which only residuals of an initial set of redexes are contracted. For that, we first extend the concept of residual to finite derivations as follows: \( A(e) := A \setminus \{ A[B;d] := (\text{Red}(B)) \} \).

Given \( A \subseteq \text{Red}(t) \), a (possibly infinite) derivation \( t \xrightarrow{\text{Red}(t)} \)

Note that our equations do not duplicate/erase/rename labels, so that any redex has a unique residual along the equivalence. Then well-definedness can be stated as the fact that every equational proof of \( t \xrightarrow{\alpha} \) induces the same bijection of redexes. The proof
of this property is based on the identification of three structural invariants of labeled redexes with respect to the equivalence \( \sim \), it is omitted for lack of space.

**Lemma 2 (Well-Definedness).** Let \( t, u \in T_{\nu \nu} \), \( t \sim u \). Then there is unique bijection \( \phi : \text{Red}(t) \to \text{Red}(u) \) s.t. \( \phi(A) = A' \) iff \( A(t \sim u)A' \).

Note that uniqueness guarantees that if \( t \triangleleft u \) and \( t \triangleright u \) are two derivation trees of \( t \sim u \), then \( A(t \triangleleft u)A' \) iff \( A(t \triangleright u)A' \).

The last result concerns preservation of residuals by means of \( \sim \)-equivalence classes, where below we write \( t \sim_\phi u \) to emphasize that \( \phi \) is the (unique) bijection given by Lemma 2.

**Lemma 3.** Let \( t \sim_\phi u \). Consider \( A, B \in \text{Red}(t) \). If \( t \xrightarrow{A} t' \), then:

1. **Simulation:** \( \exists u' \) s.t. \( u \xrightarrow{\phi(A)} u' \) and \( t \xrightarrow{\phi(A)} t' \).
2. **Same equivalence target:** \( t' \sim u' \), i.e. \( \exists \xi s.t. t' \sim_\xi u' \), and
3. **Preservation of residuals:** if \( B[A]B' \), then \( \phi(B)[\phi(A)](\phi(B)) \).

**5. Abstract Standardization**

A standard derivation is a canonical element of a class of derivations, where canonicity is expressed by means of completeness: whenever \( t \to u \) then there is a standard derivation from \( t \) to \( u \). In \( \lambda \)-calculus, a derivation is standard if redexes are selected from left-to-right. It is easy to see that the right-to-left order, instead, does not provide completeness. Consider the derivation \( t = (\lambda x. xx)(II) \to_\beta (II) \to_\beta I(II) \): it reduces both redexes in \( t \), but it can only be obtained by reducing these redexes according to a left-to-right order, as the right-to-left order would instead give \( (\lambda x. xx)(II) \to_\beta (\lambda x. xx)I \to_\beta II \). The left-to-right orientation is specific to the \( \lambda \)-calculus and may not be appropriate for other term rewriting systems, as for example the first-orderTRS containing the unique rewriting rule \( f(x,a) \to x \). The abstract theory of standardization then replaces the left-to-right order with an abstract order \( \prec \), capturing what we call the action principle.

**The Action Principle.** Let us introduce two crucial concepts. Firstly, reduction of a redex \( A \) may **directly act** on another redex \( B \) (i.e. duplicate/erase it). For example in \( \lambda \)-calculus, if \( t = (\lambda x. xx)(II) \to_\beta (II) \to_\beta I(II) \), then reduction of the leftmost redex \( A \) in \( t \) duplicates the redex \( B = II \). Secondly, reduction of a redex \( A \) may **indirectly act** on a redex \( B \) by \( 1 \) creating a new redex \( C \) that may directly (or indirectly act) on a (residual of) \( B \), or \( 2 \) changing the set of redexes that can act on \( B \) or on which \( B \) can act. An example of the first case: if \( t = (\lambda x. xx)(II)M \to_\beta M(II) \), where \( M = \lambda x. xx \), then reduction of the leftmost redex \( A \) in \( t \) indirectly acts on the redex \( B = II \) by creating the redex \( M(II) \) which can duplicate (the residual of) \( B \). An example of the second case will be given at the end of Section 8.

Standard derivations are defined abstractly according to the action principle: for every redex \( A \), reduction of \( A \) forbids future reductions of redexes which may have acted (directly or indirectly) on \( A \) before. The nesting order \( \prec \) essentially has to induce this principle by capturing as much action as possible (we will see that \( \prec \) has to capture direct action, while in general it might capture only some cases of indirect action). The abstract theory of standardization re-organizes derivations according to the action principle, relying on some axioms about the interaction between the nesting order and residuals. Then, whenever the axioms are satisfied, the standardization theorem follows: an abstract reasoning proves that \( \prec \)-ordered derivations are complete [19, 30, 32].

**Total and Partial Orders.** The left-to-right order has a fundamental property, it is a total order, i.e. given two coinitial redexes \( A \) and \( B \), then either \( A \prec B \) or \( B \prec A \). In such a case the abstract theory is quite simple.

Total orders have a limited descriptive power, though. Consider a head-normal form \( t = uv \). The left-to-right order forces the redexes in \( u \) to be reduced before those in \( v \). However, it is clear that \( u \) and \( v \) cannot interact in any way, and consequently they could be reduced in parallel even if this parallelism is not captured by the left-to-right order. In other words, the left-to-right order captures more than the action principle. Standardization can then be refined by switching to a **partial order**. For instance, the partial tree order \( \prec_{\text{tree}} \), where \( A \prec_{\text{tree}} B \) if \( B \) is a sub-term of \( A \) (so that \( A \) and \( B \) are unrelated if \( B \) is to the right of \( A \)), captures the parallelism of redexes in \( t \) and admits a standardization theorem [30, 32]. The abstract theory of standardization for partial orders is more sophisticated, so for the moment we stay with total orders and shall return to partial orders in Section 7.

In the rest of this section we introduce the necessary notions to deal with the abstract standardization theorem for total orders. In Section 6 we will adapt the left-to-right order to the \( \lambda_{\text{sub}} \)-calculus and obtain a left-to-right standardization theorem. From Section 7 to the end of the paper, we will motivate and study a relaxed variant of this order, introducing the framework for partial orders and prove a more general standardization theorem for the \( \lambda_{\text{sub}} \)-calculus.

**The Abstract Theory.** In this paper we resort to the abstract framework developed by Gonthier, Lévy, and Melliès in [19], but we adopt the more general formulation given by Melliès in his PhD thesis [30], as the axioms in [19] will not work when we shall later relax the left-to-right order. The framework is based on **Abstract Rewrite Systems** (ARS), a general formalism encompassing first and higher-order rewriting, plus a **residual relation** \( \prec \) and a **partial order** \( \prec \) on redexes, verifying some axioms. The axioms are divided into two groups: The **basic axioms**, which guarantee a well-behaved theory of residuals, and the **standardization axioms**, which concern the interaction of the order with the residual relation.

To formally define the axioms we introduce the following notion. An **Abstract Rewrite System** (ARS) is a tuple of the form \((\mathcal{O}, \mathcal{R}, \mathcal{Src}(\cdot), \mathcal{Tgt}(\cdot), \mathcal{Tgt}(\cdot)\prec_{\mathcal{T}}, \mathcal{O})\) where \( \mathcal{O} \) is called the set of objects, \( \mathcal{R} \) is called the set of redexes, \( \mathcal{Src}(\cdot) \) and \( \mathcal{Tgt}(\cdot) \) are functions from redexes to objects that we call source and target functions (resp.), \( \mathcal{O} \) is a family of binary relations indexed by the set of redexes that we call the residual relation and \( \prec_\mathcal{T} \) is a partial order on redexes. We say that two redexes \( A \) and \( B \) are **coinitial** if \( \mathcal{Src}(A) = t = \mathcal{Src}(B) \) for some term \( t \). Note that the \( \lambda_{\text{sub}} \)-calculus may be seen as an ARS where \( \mathcal{O} \) is the set of terms \( \mathcal{T} \) defined in Section 2, whereas the \( \lambda_{\text{sub}} \)-calculus is an ARS where \( \mathcal{O} \) is the set of \( \sim \)-equivalence classes generated by \( \mathcal{T} \) modulo the graphical equivalence \( \sim \). The basic axioms are:

- **Autoecrosis (AE).** For any redex \( A, \mathcal{A}(\mathcal{A}) = \emptyset \).
- **Finite residuals (FR).** Let \( A, B \) be coinitial redexes. Then the set \( \{ C \mid A(B)C \} \) is finite.
- **Uniqueness of ancestors (UA).** Let \( A, B, C \) be coinitial redexes. Then, \( B(A)B' \) and \( C(A)C' \) and \( B' = C' \) imply \( B = C \).
- **Finite Developments (FD).** Let \( A \) be a set of coinitial redexes. Then any development of \( A \) terminates.
- **Semantic Orthogonality (SO).** Let \( A, B \) be coinitial redexes. Then \( \mathcal{Tgt}(B(A)) = \mathcal{Tgt}(A(B)) \) and the relations \( \{ A \to B \} \) and \( \{ B \to A \} \) are exactly the same, where \( A, B \)
is the derivation that contracts A and then develops B(A); likewise for B; A(B) (cf. end of Sec. 3).

We consider two standardization axioms:

- **Linearity.** Let A and B be coinital redexes s.t. A \not\succ B. Then \exists B’ s.t. B(A)B’.

- **Context-freeness.** Let A, B, C be coinital redexes s.t. B(A)B’ and C(A)C’. If A \neq C then (B \sim C \iff B’ \sim C’).

The linearity axiom captures direct action: if A does not nest B then it cannot directly act (i.e. duplicate/erase) on B, and so B has exactly one residual after A. The context-freeness axiom forbids a form of indirect action of the second kind: whenever A cannot act on C (i.e. A \neq C) then it cannot grant or remove the power to act on C to any other coinital redex.

A derivation d : t \rightarrow u is said to be obtained from a derivation e : t \rightarrow u by a standardizing permutation, written e \sim d (noted d\subseteq e in [30]), if d is obtained from e by swapping two consecutive redexes which form an inversion w.r.t. \sim in d; more precisely, e \sim d if e = f; B'; \lambda; g and d = f; f; A'; h; g, and A \sim B, where A(B)A’ and h develops B(A). We also say that d is more standard than e. The classic notion of Lévy’s permutation equivalence on derivations, adapted to total orders, can be seen as the equivalence generated by standardization, indeed d : t \rightarrow u is permutation equivalent to e : t \rightarrow u, written d \equiv e, iff d\subseteq ⌈∪ \sim e⌉. In the sequel we write \equiv \sim when we wish to emphasize the underlying order t \sim.

A derivation d is standard if it is a \sim-normal form. The standardization theorem then follows by specializing Thm. 4.7 in [30] to total orders:

**THEOREM 1 (Abstract Total Standardization).** Consider any ARS equipped with a total order \sim and satisfying the basic and the standardization axioms. Then for any derivation d : t \rightarrow u there exists a standard derivation e : t \rightarrow u such that d \equiv \sim e.

The next proposition proves the semantical orthogonality property for \lambda_{sub}, that is expressed as a form of local confluence for the labeled system. Again we resort to well-labeled terms.

**PROPOSITION 2 (SO).** The \lambda_{sub}-calculus endowed with the left-to-right order satisfies Semantic Orthogonality.

**Proof.** The proof uses well-labeled terms to trace residuals in such a way that axiom (SO) can be reformulated as follows:

The reduction relations \overset{\lambda}{\rightarrow} and \overset{\beta}{\rightarrow} locally commute, i.e. if \lambda, u_1 \rightarrow u_2 \in \mathcal{T}_{\text{Red}}, \lambda \overset{\beta}{\rightarrow} u_1 and \lambda \overset{\beta}{\rightarrow} u_2 then there exists v s.t. u_1 \rightarrow v and u_2 \rightarrow v.

This alternative statement can be proved by induction on the relations \overset{\lambda}{\rightarrow} and \overset{\beta}{\rightarrow}.

**Standardization Axioms.** Totality of the order \sim provides very simple proofs of these axioms:

**PROPOSITION 3.** The \lambda_{sub}-calculus endowed with the left-to-right order satisfies Linearity and Context-freeness.

**Proof.** Linearity. By totality of \sim, we have to show that if B \sim A in t then \exists B’ s.t. B(A)B’. Now, if A is a dB-redex this is obvious, as no redex is duplicated/erased by a dB-step. If A is a \{ge, \lambda\}-redex then it can only act on redexes whose anchor is in its box, i.e. on redexes on its right, and thus not on B.

Context-freeness. If A \not\sim L C then C \sim L A. Assume B \sim L C (and so B \sim L C \sim L A). Then, A is on the right of both B and C. It is easily seen that A can only change the order between redexes on its right; consequently B’ \sim L C’. The other direction is by contraposition. Assume B \not\sim L C, that is C \sim L B. We have to prove that B’ \sim L C’, i.e. C’ \sim L B’. There are two cases. If C \sim L B \sim L A then we reason as in the previous direction, getting C’ \sim L B’. Otherwise, we have C \sim L A \sim L B. Now, the only case that is not immediate is when A is a \lambda-redex. It is enough to observe that a \lambda-step can only move the redexes in its box at most where the step itself was; hence, B’ can at most be where A was (while the position of C is left unchanged), and so C’ \sim L B’.

We can then conclude with our first standardization theorem:

**COROLLARY 1 (Left-to-right Standardization for \lambda_{sub}).** If t \rightarrow_{\lambda_{sub}} u then there is a \sim-standard \lambda_{sub}-derivation from t to u.

**Proof.** It follows from Theorem 1, whose hypothesis are given by Propositions 2, 1, and 3 (the first three basic axioms are immediately seen to hold).

7. **Towards Equational Standardization**

**Proof-Nets.** The linear substitution calculus has been designed to mimic the representation of \lambda-calculus in linear logic proof-nets [18], where \beta-reduction is decomposed into small steps. The relationship between the two formalisms occurs at the static and the dynamic levels: every term can be mapped to a proof-net, and every proof-net can be mapped to a term; and the β-reduction is decomposed into small steps. The classic notion of Lévy’s permutation equivalence on derivations, adapted to total orders, can be seen as the equivalence generated by standardization, indeed d : t \rightarrow u is permutation equivalent to e : t \rightarrow u, written d \equiv e, iff d\subseteq ⌈∪ \sim e⌉. In the sequel we write \equiv \sim when we wish to emphasize the underlying order \sim.

A derivation d is standard if it is a \sim-normal form. The standardization theorem then follows by specializing Thm. 4.7 in [30] to total orders:

**THEOREM 1 (Abstract Total Standardization).** Consider any ARS equipped with a total order \sim and satisfying the basic and the standardization axioms. Then for any derivation d : t \rightarrow u there exists a standard derivation e : t \rightarrow u such that d \equiv \sim e.

In this section we prove a standardization theorem for the \lambda_{sub}-calculus relative to a total order. This result is an extension of the left-to-right standardization theorem for \lambda-calculus.

**The Left-to-Right Order.** Given two redexes A and B in t we say that A is left-to-right nesting or on the left of B, written A \sim L B, if the anchor of A is to the left of the anchor of B (looking at t as a string of symbols). Clearly, \sim L is a total order so that A \not\sim L B and A \not\sim L B imply B \not\sim L A.

In order to prove standardization for the \lambda_{sub}-calculus by means of Theorem 1 we need to verify the basic and the standardization axioms.

**Basic Axioms.** The first three basic axioms are trivially true. By using (well-labeled terms to trace residuals we can prove the finite development axiom.

**PROPOSITION 1 (FD).** Let t \in T and let A \subseteq \text{Red}(t). Then any \lambda_{sub}-development of A terminates.

**Proof.** Consider any t \in T and lift every redex in A \subseteq \text{Red}(t) with a different label belonging to some arbitrary set L. It is clear that the resulting term is a well-labeled term. Now, let us consider the labeled reduction relation \rightarrow_{L} := \bigcup \sim L \rightarrow_{\lambda}. To prove the (FD) property it is sufficient to show that the reduction relation \rightarrow_{L} terminates on well-labeled terms. Termination is proved by means of a measure that strictly decreases with every reduction step (similar to that one used in the proof of Lemma 2 in [6]).
induces a bijection of redexes, and so it is possible to mimic derivations as follows: given a derivation \( d: t \rightarrow^* u \) and a term \( t' \rightarrow t \) we can unambiguously refer to the projection \( d': t' \rightarrow^* u' \) (with \( u' \sim u \)) of \( d \) on \( t' \). Since \( d \) and \( d' \) essentially reduce the same redexes at each step, one expects any reasonable notion of standardization to apply without distinction to both derivations in the sense that either both are standard or none of them is. This would imply, in particular, that our standardization theorem for \( \lambda_{\text{sub}} \) (Corollary 1) also applies to proof-nets. Unfortunately, the left-to-right order does not meet this requirement. Indeed, a left-to-right derivation for \( (t[x/u])[y/v] \) (where \( y \notin \text{fv}(u) \)) does not project to a left-to-right derivation for \( t[y/v][x/u] \), obtained by applying the equation \( t[x/u][y/v] \sim t[x/u][y/v] \).

**Relaxing the Order.** We are therefore going to relax the total left-to-right order \( \prec_s \) to a partial order \( \prec_b \) that will be stable by \( \sim \). The definition of \( \prec_b \) shall be guided by the action principle. 

Because of the linearity axiom, it is mandatory that \( \prec_b \) captures direct action between redexes. For that, a simple diagrammatic intuition, due to Klop [26] and then explored by Melliès [32], turns out to be extremely helpful. The idea relies on the analysis of local confluence diagrams. Whenever a redex is duplicated (resp. erased), then the standard derivation should be the longest (resp. shortest) side of the diagram. Consider for example the diagram in Fig.1-a. The standard way to get from the source to the sink of the diagram is the longest (i.e. down-bottom) side, because the 1α-redex acts on (i.e. nests) the redexes in \( s \). The standard derivation associated to the diagram in Fig.1-b—which is an erasing case—is instead the shortest side, because the erasing redex acts on the redexes in \( s \).

The previous two cases are just reformulations of similar cases in λ-calculus, and amount to what can be described as nesting as subterms. The novelty of \( \lambda_{\text{sub}} \) is that direct action (and then nesting of redexes) can also happen at a distance. Consider now the example of the diagram in Fig.1-c. The duplicated 1α-redex on \( y \) is not syntactically contained in the acting 1α-redex on \( x \). Worse, the same diagram applies to terms like \( (x[x/y][y/z])[y/z] \), where \( [x/y] \) and \( [y/z] \) are no longer next to each other. According to the action principle, our order has to impose that the 1α-redex on \( x \) nests the 1α-redex on \( y \) so that the standard side of the diagram is the longest one. Diagram in Fig.1-d is the version at a distance of the erasing diagram, requiring the same notion of nesting at a distance. All these intuitions lead to the partial order of the next section, that will be used in Section 8 to show standardization of the \( \lambda_{\text{sub}} \)-calculus.

8 Remark that otherwise one would lose completeness, as there would not be any standard derivation from \( C[x][x/s] \) to the term \( C[[s]][x/s] \), which is the intermediate term in the sequence \( C[x][x/s] \rightarrow C[[s]][x/s] \).

**8. The Box Standardization Theorem.** We now generalize the notions introduced in Section 5 to the partial box order, and then prove the standardization theorem for the \( \lambda_{\text{sub}} \)-calculus with respect to this order.

**The Box Order.** Let \( A, B \in \text{Red}(l) \). Then,

- **A immediately boxes**, \( B \), noted \( A \prec^+_B \), if the anchor of \( B \) (i.e. the variable possibly carrying a label) is in the box of \( A \). i.e. if the pattern of \( A \) is any of \( (\lambda x.B)u \), \( C[x][x/u] \) or \( t[x/u] \), then the anchor of \( B \) appears in \( u \).
- **A boxes B**, noted \( A \sim B \), if \( A(\sim)B \) (we use \( A \sim B \) for \( A(\sim)B \)).
- **A and B are disjoint**, noted \( A \parallel B \), if \( A \nmid B \) and \( B \nmid A \).

Note that \( A \prec_b B \) implies \( A \prec_l B \). Additionally, note the transitive closure in the definition of \( \prec_b \)—unnecessary for \( \prec_l \)—which impacts on the proofs of the standardization axioms for \( \prec_b \).

The key property of the box order is that it is stable by the equivalence \( \sim \). For example, for \( t[x/u][y/v] \) with \( y \notin \text{fv}(u) \) the redexes in \( u \) and the redexes in \( v \) are not related by \( \prec_b \), so that \( \sim \) is stable by the permuting axiom \( t[y/v][x/u] \sim t[x/u][y/v] \) (where \( y \notin \text{fv}(u) \) and \( x \notin \text{fv}(v) \)). More precisely, given \( s \sim t \), the bijection between \( \text{Red}(s) \) and \( \text{Red}(t) \) defined in Section 4 is order-preserving. To show this property it is sufficient to remark that symbols inside boxes never go in/out these boxes by means of the equivalence relation. Thus, we get:

**Lemma 4 (Preservation of the Box Order by Equivalence).** Let \( t, u \in T \) s.t. \( t \sim u \), \( \phi \) is the bijection specified in Lemma 2. Then, \( \phi \) commutes with \( \prec_b \), i.e. \( A \prec_b B \) iff \( \phi(A) \prec_b \phi(B) \).

Several remarks on \( \prec_b \) are in order:

1. **Disjoint redexes may superpose:** \( A \parallel B \) does not necessarily imply that \( A \) and \( B \) are syntactically disjoint. Examples: the redexes \( A_\alpha \) and \( B_\beta \) are disjoint but 1) syntactically superposed in \( (x^x\beta)[x/y] \) and 2) syntactically nested in \( (\lambda x.x^x)[x/z]y \). However, disjoint redexes always strongly locally commute in the following sense: if \( t_0 \vdash x \rightarrow t_1 \) and \( t_1 \vdash x \rightarrow t_2 \) then there exists \( t_3 \) s.t. \( t_1 \rightarrow x \vdash t_3 \) and \( t_2 \vdash t_3 \rightarrow x \vdash t_3 \). Note that this is just a particular case of axiom (SO) where the diagram can be closed by using just one reduction step from \( t_1 \) to \( t_3 \).

2. **Indirect action:** the box order fails to capture indirect action, in the sense that \( A \parallel B \) implies that \( A \) cannot directly act on \( B \) but it can still indirectly act on it (and viceversa). For example, in the following derivation \( A_\alpha \parallel B_\beta \) and \( A_\alpha \) creates a redex that can act on \( B_\beta \):

\[
x^x[x/\lambda y.y][((\lambda z^z).z)u] \rightarrow_{1a} ((\lambda y.y)[x/\lambda y.y][((\lambda z^z).z)u])
\]

This apparently odd fact—that is not specific to the box order—will not forbid the standardization theorem. Its consequences, and an easy way to deal with them, will be discussed in Section 10. Note that by definition the box order is the minimum standardization order, in the sense that it is the transitive closure of the relation capturing direct action only.

3. **The order on \( \lambda \)-terms:** in [32] (pp. 74-75) Melliès considers three orders on \( \lambda \)-terms, the left-to-right order plus two refinements called the tree and the argument order. When restricted to \( \lambda \)-terms, our order coincides with the argument order.

4. **Linear logic interpretation:** we explained that there are two kinds of nesting in \( \lambda_{\text{sub}} \), namely, subterm nesting and nesting at a distance. In the definition of the box order they are both captured by the notion of immediately boxes. In terms of proof-nets, however, they have different interpretations, despite they are both related to the concept of exponential box 7.

**Standardization Up to Square Equivalence.** A consequence of switching to partial orders is that the notion of standard derivation has to be refined. In fact, partial orders force to work modulo the exchanges of disjoint redexes, even to simply define what is an ordered (i.e. standard) sequence; this is why abstract standardization is a form of rewriting modulo.

Standardizing permutations are defined as in Section 5. However, to deal with disjoint redexes and adequately represent their

7 A subterm nests \( B \) when \( B \) is contained in the !-box of \( A \), while \( A \) nests \( B \) at a distance when the cut corresponding to \( B \) crosses an auxiliary conclusion of the box of \( A \).
parallelism induced by the partial order, a new concept is necessary. A derivation \( d \) is said to be obtained from a derivation \( e \) by square permutation, written \( d \overset{\triangle}{\sim} e \), if \( d = f; A; B \); and \( e = f; B; A \); and \( A \parallel B \), where \( A(B)A' \) and \( B(A)B' \). We write \( \triangle \) for the equivalence relation generated by \( \overset{\triangle}{\sim} \), and call it square equivalence. We use \( \sim_{\triangle} \) for the relation \( \sim_{\triangle} \) modulo \( \triangle \) and we generalize the permutation equivalence relation \( \sim_{\triangle} \) as \( (\sim_{\triangle} \cup \sim_{\triangle}^{-1})^* \).

A derivation \( d \) is standard if it is a \( \sim_{\triangle} \)-normal form modulo \( \triangle \). The standardization theorem (Thm. 4.7 in [30]) then follows:

**Theorem 2 (Abstract Partial Standardization).** Consider any ARS equipped with a partial order that enjoys the basic and the standardization axioms. Then for every derivation \( d \) there exists a standard derivation \( e \) s.t. \( d \sim_{\triangle} e \).

**Results.** Our goal is to prove a standardization theorem for \( \lambda_{\text{sub}} \)-calculus endowed with the box order \( \prec_{\text{box}} \). The proof consists in first obtaining the theorem for \( \lambda_{\text{sub}} \) and then lifting it to \( \lambda_{\text{sub}} \).

The basic axioms do not mention the order, and thus they still hold for \( \lambda_{\text{sub}} \). The standardization axioms are proved by the following proposition:

**Proposition 4.** The \( \lambda_{\text{sub}} \)-calculus endowed with the box order \( \prec_{\text{box}} \) satisfies Linearity and Context-freeness.

**Proof.** Linearity is easy, by just remarking that a redex \( B \) can be duplicated or erased by \( A \) only if its anchor is in the box of \( A \), case in which \( A \prec_{\text{box}} B \). Context-freeness requires a detailed study of the interaction between the box order \( \prec_{\text{box}} \) and the reduction relation. In particular the proof is involved due to the transitive clause defining the box order, and it is omitted for lack of space.

Then, the previous proposition and Theorem 2 give:

**Corollary 2 (Box Standardization for \( \lambda_{\text{sub}} \)).** If \( t \prec_{\lambda_{\text{sub}}} u \) then there exists a \( \prec_{\text{box}} \)-standard \( \lambda_{\text{sub}} \)-derivation from \( t \) to \( u \).

In order to lift the previous result to \( \lambda_{\text{sub}} \) we use the results on equivalence classes developed in Section 4.

**Proposition 5.** The \( \lambda_{\text{sub}} \)-calculus endowed with the box order \( \prec_{\text{box}} \) satisfies the basic and the standardization axioms.

**Proof.** By Lemma 3 and Lemma 4 the notion of residual and the box order lift to \( \sim_{\text{box}} \)-equivalence classes preserving their properties. Moreover, the FD axiom can be shown exactly as done for \( \lambda_{\text{sub}} \) in Proposition 1, by means of a measure that strictly decreases with every reduction step and remains equal for every pair of equivalent pairs. Therefore, all axioms hold for \( \lambda_{\text{sub}} \) and we can conclude.

Proposition 5 allows to apply Theorem 2 again and finally obtain our nonstandard, equational standardization theorem.

**Corollary 3 (Box Standardization for \( \lambda_{\text{sub}} \)).** If \( t \prec_{\lambda_{\text{sub}}} u \) then there is a \( \prec_{\text{box}} \)-standard \( \lambda_{\text{sub}} \)-derivation from \( t \) to \( u \).

The next section will strengthen the result by showing uniqueness (modulo \( \triangle \)) of the obtained standard derivation. We conclude this section with a remark.

---

**Figure 1.** Some standardization diagrams.

<table>
<thead>
<tr>
<th>Redex Order</th>
<th>Existence</th>
<th>Uniqueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{\text{sub}} )</td>
<td>Total, left-to-right (Sec. 6)</td>
<td>Thm. 1</td>
</tr>
<tr>
<td>( \lambda_{\text{sub}} )</td>
<td>Partial, box (Sec. 8)</td>
<td>Cor. 2</td>
</tr>
<tr>
<td>( \lambda_{\text{sub}} )</td>
<td>Partial, box (Sec. 8)</td>
<td>Cor. 3</td>
</tr>
</tbody>
</table>

**Figure 2.** Summary of results on standardization

**Variation on the Context-freeness Axiom.** We use in this paper Context-freeness as it appears in Melliès’s thesis [30]. In GLM’s theory [19], however, the axiom appears in a stronger, symmetric form which does not hold for \( \lambda_{\text{sub}} \). Indeed,

- **Strong context-freeness.** Let \( A, B, C \) be coinitial redexes s.t. \( B(A)B' \) and \( C(A)C' \). If \( A \neq C \) then \((B = C \Leftrightarrow B' \prec C')\) and \((C \prec B \Leftrightarrow C' \prec B')\).

Consider \( t = (\lambda x^a.y[x/y])(z^u/x)\), where \( A_x \parallel C_y, C' \parallel B_{z'}, \) and \( A_{\alpha} \prec_{\text{box}} B_{\beta} \). The reduction step of the redex \( A_{\alpha} \) yields the term \( y^{g}[x/z]((z^{u}/x)\parallel z/u)\), where \( C' \prec_{\text{box}} B_{\beta} \), contradicting the strong form of the axiom (but not the weaker one we use) and providing an example of the second kind of indirect action, as \( A_{\alpha} \) which can act on \( (\text{the residual of}) \) \( B_{\beta} \).

**9. Uniqueness**

The standardization results obtained up to now concern the existence of a standard derivation. In this section the theorems are strengthened by proving uniqueness of such derivations. We first discuss the simple case of the left-to-right order, and then provide an original proof of uniqueness (modulo) for the box order.

The axiomatic theory in Melliès’s thesis provides uniqueness of standard derivations when two further axioms, Enclave and Stability, hold (Theorem 4.5 in [30]).

Axiom Enclave has two parts. Let \( A, B, B' \) be coinidal redexes s.t. \( B(A)B' \).

1. **Creation.** If \( B \prec A \) and \( A \) creates a redex \( C \), then \( B' \prec C \).
2. **Nesting.** Let \( C, C' \) be redexes. If \( B \prec A \prec C \) and \( C(A)C' \), then \( B' \prec C' \).

In terms of the action principle Enclave forbids yet another form of indirect action: if \( B \) can directly act on \( A \) then \( A \) cannot indirectly act on \( B \). Consider the creation part: if \( C \prec B' \) or \( C \parallel B' \) then \( C \) may act on \( B' \), and so \( A \) would indirectly act on \( B \), similarly for the nesting part.

**Proposition 6.** The \( \lambda_{\text{sub}} \)-calculus endowed with the left-to-right order \( \prec_{L} \) satisfies Enclave.

**Proof.** Creation: a simple inspection of the cases of creation (Section 3) shows that a redex \( A \) can create a redex \( C \) only on its right or at most where it was, so that \( C \) cannot nest any redex that was on the left of \( A \). Nesting: similarly, it is easily seen that \( A \) can only move redexes on its right and at most where it was.

In the case of a total standardization order the axiom Stability disappears, because its statement (see [30]) assumes two \( \prec_{\triangle} \)-disjoint redexes. Thus, by applying Melliès’ theorem we obtain:
Theorem 3 (Uniqueness for $\lambda_{\text{sub}}$). If $t \xrightarrow{\text{L}} u$ then there exists a unique $\prec_{s}$-standard $\lambda_{\text{sub}}$-derivation from $t$ to $u$.

Switching to partial orders, uniqueness is necessarily relaxed to uniqueness modulo $\sim$ (as standard derivations are defined up to $\sim$).

Interestingly, for the partial box order both points of Enclave fail for $\lambda_{\text{sub}}$. For creation consider the step:

$$s = ((\lambda x^a.x)(y[e^a/z^a]))[z^a/u] \xrightarrow{\Delta} ((\lambda z^a.x)y)[z^a/u] = t$$

Let us call $A_{e}$ and $B_{e}$ the two labeled redexes in $s$, and $B_{e}$ (resp. $C$) the residual of $B_{e}$ (resp. the created ge-redex on $z^a$) in $t$. We have $B_{e} \prec_{s} A_{e}$ but $B_{e} \not\prec_{s} C$. For nesting it is enough to consider a redex $C$ inside $u$ in the counter-example for creation.

So the box order does not capture the indirect action expressed by Enclave, and Mellies’ axiomatics cannot provide uniqueness modulo $\prec_{s}$ for $\lambda_{\text{sub}}$. However, we can obtain it via another argument. In particular, our alternative proof shows that Enclave (together with Stability) is a sufficient but by no means necessary condition for uniqueness. We first show uniqueness modulo $\prec$ for $\lambda_{\text{sub}}$, and then lift this result to $\lambda_{\text{sub}}$.

Theorem 4 (Uniqueness Modulo for $\lambda_{\text{sub}}$). If $t \xrightarrow{\text{L}} u$ then there exists a $\prec_{s}$-standard $\lambda_{\text{sub}}$-derivation from $t$ to $u$ that is unique modulo $\prec$.

Proof. (Sketch) Let $e$ be the $\prec_{s}$-standard $\lambda_{\text{sub}}$-derivation given by Corollary 2 and $f$ be the unique $\prec_{s}$-standard derivation given by Corollary 1. We prove that $e \sim_{s} f$, from which the statement follows. Consider the leftmost contracted redex $C$ in $e = B_{1}; \ldots; B_{n}$ (i.e. the leftmost redex in $\{ A \in \mathbb{R}(t) \mid A(B_{1}; \ldots; B_{i})B_{i+1}, 0 \leq i < n \}$) and let $e'$ be $C; e_{1}; e_{2}$ where $e_{1} = (B_{1}; \ldots; B_{i})C$ and $C(B_{1}; \ldots; B_{i})B_{i+1}$ and $e_{2} = B_{i+1}; \ldots; B_{n}$. It is not difficult to prove, given that $e$ is $\prec_{s}$-standard, that $e' \sim_{s} e$, $|e_{1}| = 1$ and so $|e_{1}; e_{2}| = |e| - 1$, where $|e|$ denotes the number of steps in $e$. By induction hypothesis we obtain a $\prec_{s}$-standard derivation $f'$ s.t. $f' \sim_{s} e_{1}; e_{2}$. Therefore $C; f'$ is also $\prec_{s}$-standard and moreover $(C; f') \sim_{s} (C; e_{1}; e_{2}) \sim_{s} e$. Thm. 3 implies $C; f' = f$.

Note that the proof does not depend concretely on $\lambda_{\text{sub}}$, but only on the fact that $\prec_{s}$ can be embedded in a total order admitting a standardization theorem, which is a fully abstract argument.

By exploiting again the results on equivalence classes, we obtain the strongest result of the paper:

Theorem 5 (Uniqueness Modulo for $\lambda_{\text{sub}}$). If $t \xrightarrow{\text{L}} u$ then there exists a $\prec_{s}$-standard $\lambda_{\text{sub}}$-derivation from $t$ to $u$ that is unique modulo $\sim$.

Proof. By Lemma 3 and Theorem 4 using the stability of the $\prec_{s}$-order by the equivalence $\sim_{s}$ given by Lemma 4.

10. External Redexes, Coinductively

When the nesting order $\prec$ is total, standardization is relatively easy, because a derivation $d : t \rightarrow u$ may be standardized by simply selecting the minimum redex $A$ among the redexes in $t$ that are reduced in $d$ and then repeating the process for the residual of this derivation after $A$. However, when the order is partial—as our box order, and more generally in Gonthier-Levy-Melliès’ approach—there may be many $\prec$-minimal redexes among the redexes in $t$ that are reduced in $d$; and randomly selecting one of them does not necessarily produce a standard derivation. This is due to the fact that the partial box order—as it is usually the case with partial orders—fails to properly capture indirect action. Indeed, even if two minimal redexes $A$ and $B$ are necessarily disjoint, $A$ may create a redex which will nest (a residual of) $B$.

We illustrate this situation in the simpler setting of the $\lambda$-calculus using the box order (called argument order by Mellies in [32]). Let $I^{\alpha}$ stand for $\lambda x^{a}.x$ and $A_{\alpha}$ for $I^{\alpha} I$ (and similarly for $B_{\beta}$). Consider the derivation:

$$(I^{\alpha} I) (I^{\beta} I) \xrightarrow{\beta_{\gamma}} (I^{\alpha} I) \xrightarrow{\beta_{\gamma}} I \rightarrow_{\beta} I$$

Note that both $A_{\alpha}$ and $B_{\beta}$ are minimal in the initial term, and the derivation picks at every step a minimal redex. However, the derivation $d$ is not standard. This can be seen by permuting $A_{\alpha}$ and $B_{\beta}$, which yields the following $\sim$-equivalent derivation, where the two last steps form an inversion for the box order:

$$(I^{\alpha} I) (I^{\beta} I) \xrightarrow{\beta_{\gamma}} I (I^{\beta} I) \xrightarrow{\beta_{\gamma}} (I^{\beta} I) \xrightarrow{\beta_{\gamma}} I$$

A standardizing permutation swapping them produces the following standard derivation:

$$(I^{\alpha} I) (I^{\beta} I) \xrightarrow{\beta_{\gamma}} I (I^{\beta} I) \xrightarrow{\beta_{\gamma}} (I^{\beta} I) \xrightarrow{\beta_{\gamma}} I$$

This example shows that selecting minimal redexes does not necessarily give standard derivations when the order is partial. A similar example in $\lambda_{\text{sub}}$ is obtained by considering the derivations of $x^{a}[x]\lambda z y((\lambda z^{b}.z)u)$. The solution to this problem is to select an external redex [8, 10, 19, 22, 28, 30, 38, 39], i.e. a minimal one on which no other redex can indirectly act. The definition of external redex for a derivation in GLM’s theory is given via an extraction process. In particular, the definition mentions the nesting order only indirectly. Alternatively, external redexes can be defined coinductively as persistently minimal redexes, as hinted in [38], i.e. the minimal redexes whose (unique) residual after any other redex is still persistently minimal.

We first recall the definition of external redex from [30], that requires two preliminary definitions. A redex $A$ traverses a coinduction derivation $d$ becoming $B$, written $A \xrightarrow{d} B$, iff:

• For $d = e$, $A \xrightarrow{e} B$ iff $B = A$.

• For $d = C; e$, $A \xrightarrow{C; e} B$ iff $C \not\xrightarrow{e} A'$ and $A' \xrightarrow{e} B$, where $A(C)A'$.

A redex $A$ in $d$ can be extracted from a derivation $d : t \rightarrow u$, written $A \xrightarrow{d} d$, if there exist $d_{1}, d_{2}$ s.t. $d = d_{1}; A'; d_{2}$ and $A \xrightarrow{d_{2}} A'$. A redex $A$ in $d$ is external for $d : t \rightarrow u$ if $A \in \text{Ext}_{d}(d) := \{ B \mid \forall e \equiv e. d . B \not\xrightarrow{e} \}$. The following result corresponds to Lemma 4.36 in [30].

Lemma 5 (External Gives Standard). Consider any ARS equipped with a partial order $\prec$ that enjoys the basic and the standardization axioms. If $d = A_{1}; \ldots; A_{n}$ and $A_{i} \in \text{Ext}_{d}(A_{i}; \ldots; A_{n})$ ($1 \leq i \leq n$), then $d$ is $\prec$-standard.

External redexes generalize leftmost redexes in $\lambda$-calculus, whose key properties are 1) no other redex can act on it, and 2) its unique residual after any other redex is still the leftmost redex. This suggests to define external redexes coinductively as follows.

Let $d : t \rightarrow u$ be a derivation. The set of starting redexes of $d$ is $d^{\prime} := \{ B \mid \exists e \text{ s.t. d } \equiv_{e} B ; e \}$. A redex $A$ in $d$ is $\prec$-external for $d$ if $d$ is not empty and:

1. Minimality: $A$ is minimal in $d^{\prime}$, and

2. Persistency: whenever $d \equiv_{e} B; e$ and $A(B)A'$ ($A'$ is unique by minimality), then $A'$ is $\prec$-external in $e$.

The next two technical lemmas are used to relate the two notions of externality. Their proofs are easy:

Lemma 6. Let $A$ be $\prec$-external for $d$. Then:

1. $A \not\xrightarrow{d}$, and

2. If $e \equiv_{e} d$ then $e^{1} = d^{1}$ and $A$ is $\prec$-external for $e$.

Lemma 7. Let $A \not\xrightarrow{d}$, i.e. $d = e; A'; f$ and $A \xrightarrow{e} A'$. Then:
A for external with respect to any reduction step (and thus wrt any redex, is normalizing for λ derivation), when the following holds:

We conclude the section with the equivalence of the two notions of external redex.

**Proposition 7.** A is ≺-external for d iff A ∈ Ext≺(d).

Proof. (⇒) Let e ≺ e. Lemma 6.2 implies that A is ≺-external for e, and Lemma 6.1 gives A ≺ e, i.e. A ∈ Ext≺(d), ≺e) By coinduction on the definition of ≺-external. A ∈ Ext≺(d) implies A, and thus A is ≺-external for d.

**11. Applications**

In this section we apply the new characterization of external redexes to obtain a normalization theorem, and to prove that (weak) linear head reduction is standard.

**A Normalizing Strategy for λab and λab**

Using the coinductive reformulation of external redexes, we can now provide an easy proof that the leftmost strategy, which always reduces the leftmost redex, is normalizing for λ proof that the leftmost strategy, which always reduces the leftmost redex is standard.

**Theorem 6 (Normalization).**

1. The leftmost redex is universally ≺-external λab.

2. The leftmost strategy is normalizing for λab and λb.

Proof. 1) By coinduction. Left-to-right order: the leftmost redex is ≺-minimum and by the Enclace axiom for ≺ (Proposition 6) its residual after any other redex is still minimum and thus leftmost. We conclude using the coinductive hypothesis. 2) For λab it follows by Theorem 5.2 in Mellies’ thesis [30] (p. 137, formulating the partial box order on redexes, as we now explain. A crucial point of our definition is that that →H—in contrast to →W—is nondeterministic, for instance we have both (λx.y[y/w])z →H (λx.w[y/w])z and (λx.y[y/w])z →H y[y/w][x/z]. This fact is not a drawback, rather a plus, as it faithfully mimics the parallelism of cut-elimination in proof-nets. A simple case analysis shows that LHR enjoys the diamond property, i.e. any two steps commute and no duplication/erasures is involved. Therefore, the nondeterminism is harmless. In particular, all maximal LHR derivations (if any) have the same length. Concerning standardization, however, not all LHR derivations are standard with respect to the left-to-right order ≺, consider:

\[(λx.y[y/w])z →H (λx.w[y/w])z →H w[y/w][x/z]\]

This fact is disappointing because the nondeterminism of LHR is only apparent, and so one would like to consider LHR sequences as standard, without having to re-organize them. The box order turns LHR into a standard strategy, giving to the presentation of LHR with explicit substitutions a solid status.

**Theorem 7 (Linear Head Reduction is Standard).**

1. Linear head redexes are universally ≺-external.

2. Linear head reduction is ≺-standard.

Proof. 1) By coinduction on the definition of universally ≺-external. Universal minimality: the anchor of →H-redexes is out of all boxes, so they can never be nested by another redex. Universal persistency: it is easily seen that no redex can move the anchor of a →H-redex inside a box. So the residual of a →H-redex is a →H-redex and we conclude by the coinductive hypothesis. 2) A

\(\text{ The one reducing only cuts out of all } t!\text{-boxes that do not involve the auxiliary conclusion of any box.} \)

**Proposition 8 (Weak Linear Head Reduction is Standard).**

1. Every term t ∈ T has at most one →H redex, and if it does then it is the leftmost redex of t.

2. →H derivations are ≺-standard (and thus also ≺-standard).

Proof. 1) By induction on t. 2) By Lemma 5 and Theorem 6.
universally external redex is easily seen to be external for every coinital derivation. Then we conclude applying Lemma 5.

12. Conclusions

We study standardization for the linear substitution calculus, a calculus with explicit substitutions, that is not syntactically orthogonal and that it is equipped with an equational theory that makes it isomorphic to linear logic proof-nets. Our main result is a standardization theorem, nonstandard because it is based on a partial rather than a total order, it lifts to equivalence classes (i.e. proof-nets), and it provides a notion of standard derivation which departs from the one by levels in the linear logic literature.

Along the way, we provided other results: 1) a theory of residuals lifting to equivalence classes; 2) a simple left-to-right standardization theorem; 3) a coinductive characterization of the notion of external redex; 4) a normalization theorem; 5) a simple proof that (weak) linear head reduction is standard.

Last, we believe that a further contribution of this paper is the final analysis of Gonthier, Lévy, and Melliès’ axiomatic framework, which gives new intuitions on their complex notions.

Acknowledgments

We are grateful to Paul-André Melliès for fruitful discussions. This work was partially funded by the French-Argentinian Laboratory in Computer Science INFINIS (www.infinis.org), by the French-Argentinian project ECOS-Sud and by the Qatar National Research Fund under grant NPRF 09-1107-1-168.

References