A Linear Decomposition of Multiparty Sessions for Safe Distributed Programming

Alceste Scalas¹, Ornela Dardha², Raymond Hu³, and Nobuko Yoshida⁴

¹ Imperial College London, UK
alceste.scalas@imperial.ac.uk
² University of Glasgow, UK
ornela.dardha@glasgow.ac.uk
³ Imperial College London, UK
raymond.hu@imperial.ac.uk
⁴ Imperial College London, UK
n.yoshida@imperial.ac.uk

Abstract
Multiparty Session Types (MPST) is a typing discipline for message-passing distributed processes that can ensure properties such as absence of communication errors and deadlocks, and protocol conformance. Can MPST provide a theoretical foundation for concurrent and distributed programming in “mainstream” languages? We address this problem by (1) developing the first encoding of a full-fledged multiparty session π-calculus into linear π-calculus, and (2) using the encoding as the foundation of a practical toolchain for safe multiparty programming in Scala.

Our encoding is type-preserving and operationally sound and complete. Crucially, it keeps the distributed choreographic nature of MPST, illuminating that the safety properties of multiparty sessions can be precisely represented with a decomposition into binary linear channels. Previous works have only studied the relation between (limited) multiparty and binary sessions via centralised orchestration means. We exploit these results to implement an automated generation of Scala APIs for multiparty sessions, abstracting existing libraries for binary communication channels. This allows multiparty systems to be safely implemented over binary message transports, as commonly found in practice. Our implementation is the first to support distributed multiparty delegation: our encoding yields it for free, via existing mechanisms for binary delegation.

1998 ACM Subject Classification D.1.3 Concurrent Programming; D.3.1 Formal Definitions and Theory; F.3.3 Studies of Program Constructs — Type structure

Keywords and phrases process calculi, session types, concurrent programming, Scala

Digital Object Identifier 10.4230/LIPIcs.ECOOP.2017.24

Supplementary Material ECOOP Artifact Evaluation approved artifact available at http://dx.doi.org/10.4230/DARTS.3.2.3

1 Introduction
Correct design and implementation of concurrent and distributed applications is notoriously difficult. Programmers must confront challenges involving protocol conformance (are messages

* Partially supported by EPSRC (grants EP/K034443/1, EP/K011715/1, EP/L00058X/1, EP/N027833/1, EP/N028201/1) and EU (FP7 612985 “Upscale”). Dardha was awarded a SICSA PECE bursary for visiting Imperial College London in January–March 2016.

© Alceste Scalas, Ornela Dardha, Raymond Hu, and Nobuko Yoshida; licensed under Creative Commons License CC-BY
Editor: Peter Müller; Article No. 24, pp. 24:1–24:31
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
sent/received according to a specification?) and communication mechanics (how are the interactions actually performed?). These difficulties are worsened by the potential complexity of interactions among multiple participants, and if the communication topology is not fixed.

For example, consider a common scenario for a peer-to-peer multiplayer game: the clients, initially unknown to each other, connect to a “matchmaking” server, whose task is to group players and set up a game session in which they can interact directly. Figure 1 depicts this scenario: Q is the server, connected to three clients \( P_a \), \( P_b \), and \( P_c \). To set up a game, \( Q \) sends to each client some networking information (denoted by \( s[a]/s[b]/s[c] \), payloads of the PlayA/B/C messages) to “introduce” the clients to each other and allow them to communicate. Then, the clients follow the game protocol (marked as “Game”), consisting in some initial message exchanges (Info), and a game loop: \( P_a \) chooses a message to send to \( P_b \) (Mov1AB or Mov2AB) followed by a message from \( P_b \) to \( P_c \), who chooses which message send back to \( P_a \).

Figure 1 features structured protocols with inter-role message dependencies, and a dynamic communication topology (starting client-to-server, becoming client-to-client). Implementing them is not easy: programmers would benefit from tools to statically detect protocol violations in source code, and realise the communication topology changes.

Multiparty Session Types (MPST) [27] are a theoretical framework for channel-based communication, capable of modelling our example. In MPST, participants are modelled as roles (e.g., game players \( a, b, c \)) and programs are session \( \pi \)-calculus processes; the “networking information payloads” \( s[a]/s[b]/s[c] \) can be modelled as multiparty channels, for interpreting roles \( a/b/c \) on the game session \( s \). Notably, channels can themselves be sent/received: this allows to delegate a multiparty interaction to another process, thus changing the communicating topology. In Figure 1, the server \( Q \) sends (i.e., delegates) the channel \( s[b] \) to \( P_b \); the latter can then use \( s[b] \) to interact with the processes owning channels \( s[a] \) and \( s[c] \) (i.e., \( P_a \) and \( P_c \), after two more delegations).
The MPST framework formalises protocols as session types: structured sequences of inputs/outputs and choices. The MPST typing system assigns such types to channels, and checks the processes using them. In our example, channel \( s[b] \) could have type:

\[
S_b = \text{c!InfoBC(String). a?!InfoAB(String).} \\
\mu t. (a \& \{ \text{?Mov1AB(Int). c!Mov1BC(Int).}. \text{, ?Mov2AB(Int). c!Mov2BC(Int).}. t ) \}
\]

\( S_b \) says that \( s[b] \) must be used to realise the Game interactions of \( P_b \) in Figure 1: first to send InfoBC(String) to c, then receive InfoAB from a, then enter the recursive game “loop” \( \mu t.(...) \). Inside the recursion, a & ( ...) is a branching from a: depending on a’s choice, the channel will deliver either Mov1AB(Int) (in which case, it must be used to send Mov1BC(Int) to c, and loop), or Mov2AB (then, it must be used to send Mov2BC to c, and loop). Analogous types can be assigned to \( s[a] \) and \( s[c] \). Delegation is represented by types like \( q?PlayB(S_b) \). end, meaning: from role \( q \), receive a message PlayB carrying a channel that must be used according to \( S_b \) above; then, end the session. Session type checking ensures that, e.g., process \( P_b \) uses its channels abiding by the types above — thus safely implementing the expected channel dynamics and fulfilling role \( b \) in the game. Finally, MPST can formalise the whole Game protocol in Figure 1 as a global type, and validate that it is deadlock-free; then, via typing, ensure that a set of processes interacts according to the global type (and is, thus, deadlock-free).

**MPST in practice: challenges.** MPST could offer a promising formal foundation for safe distributed programming, helping to develop type-safe and deadlock-free concurrent programs. However, bridging the gap between theory and implementation raises several challenges:

- **C1** Multiparty sessions can have 2, 3 or more interacting roles; but in practice, communication occurs over binary channels (e.g., TCP sockets). Can multiparty channels be implemented as compositions of binary channels, preserving their type safety properties?
- **C2** MPST are far from the types of “mainstream” programming languages, as shown by \( S_b \) above. Can they be rendered, e.g., as objects? If so, what are their API and internals?
- **C3** How should multiparty delegation be realised, especially in distributed settings?

The current state-of-the-art has not addressed these challenges. On one hand, existing theoretical works on encoding multiparty sessions into binary sessions [8, 9] introduce centralised medium (or arbiter) processes to orchestrate the interactions between the multiparty session roles: hence, they depart from the choreographic (i.e., decentralised) nature of the MPST framework [27], and preclude examples like our peer-to-peer game in Figure 1. On the other hand, there are no existing implementations of full-fledged MPST; e.g., [57, 32, 33, 42, 52, 61, 55] only support binary sessions, while none of [29, 64, 17, 20] support session delegation.

**Our approach.** In this work, we tackle the three challenges above with a two-step strategy:

- **S1** we give the first choreographic encoding of a “full” MPST calculus into linear π-calculus;
- **S2** we implement a multiparty session API generation for Scala, based on our encoding.

By step **S1**, we formally address challenge **C1**. Linear π-calculus provides a theoretical framework with typed channels that cater only for binary communication, and may only be used once for input/output. These “limitations” are key to the practicality of our approach. In fact, they force us to figure out whether multiparty channels can be represented by a decomposition into binary channels — and whether multiparty session types can be represented by a decomposition into linear types. To solve these issues, we need study how to “decompose” the intricate MPST theory in (much simpler) π-calculus terms. This endeavour was not tackled before, and its feasibility was unclear. Its practical payoff is that linear π-calculus
channels/types are amenable for an (almost) direct object-based representation (shown in [61]): this tackles challenge C2. Further, using π-calculus we can prove whether such a decomposition is “correct”, i.e., whether MPST processes can be encoded to only interact on binary channels, preserving their type-safety and behaviour and “inherit” deadlock-freedom.

In step S2, we generate high-level typed APIs for multiparty session programming, ensuring their “correctness” by reflecting the types and process behaviours formalised in step S1. Following the binary decomposition in step S1, we can implement such APIs as a layer over existing libraries for binary sessions (available for Java [30], Haskell [57, 32, 42], Links [44], Rust [33], Scala [61], ML [55]), in a way that solves challenge C3 “for free”.

Contributions. We present the first encoding (Section 5) of a full multiparty session π-calculus (Section 2) into standard π-calculus with linear, labelled tuple and variant types (Section 3).

- We present a novel, streamlined MPST formulation, sharply separating global/local typing. Using this formulation, we “close the gaps” between the intricacies of the MPST theory and the (much simpler) π-calculus, and spot a longstanding issue with type merging [18] (Definition 2.9, Section 2.1 “On Consistency”). We fix it, with a revised subject reduction (Theorem 2.16).

- At the heart of our encoding there is the discovery that the type safety property of MPST is precisely characterised as a decomposition into linear π-calculus types (Theorem 6.3).

- We subsume the encodings of binary sessions into π-calculus [14, 15], and support recursion (Section 4), which was not properly handled in [13]. Further, we show that multiparty sessions can be encoded into binary sessions choreographically, i.e., while preserving process distribution (homomorphically w.r.t. parallel composition), in contrast to [8, 9].

In Section 7, we use our encoding as formal basis for the first implementation of multiparty sessions supporting distributed multiparty delegation, over existing Scala libraries (paper’s artifact1).

Conventions. Derivations use single/double lines for inductive/coinductive rules. Recursive types μt.T are always closed, and guarded: e.g., μt1,...,μtn.t, is not a type. We define unf(μt.T) = unf(T[μt.T/τ]), and unf(T) = T if T ≠ μt.T’. Type equality is syntactic: μt.T is not equal to unf(μt.T). We write P→P’ for process reductions, →∗ for the reflexive+transitive closure of →, and P→ P’ iff ∃P’ such that P→P’. We assume a basic subtyping ≤B capturing e.g. Int ≤B Real. For readability, we use blue/red for multiparty/standard π-calculus.

2 Multiparty Session π-Calculus

In this section we illustrate a multiparty session π-calculus [27] (Definition 2.1), and its typing system — including recursion, subtyping [19] and type merging [67, 18] (Section 2.1). The calculus models processes that interact via multiparty channels connecting two or more participants: this is a departure from many “classic” and simpler process calculi, like the

1 http://dx.doi.org/10.4230/DARTS.3.2.3
linear π-calculus (Section 3), that model binary channels. We provide various examples based on the scenario in Section 1.

**Definition 2.1.** The syntax of multiparty session π-calculus processes and values is:

- Processes: \( P, Q ::= 0 | P \mid Q | (\nu s)P \) (inaction, composition, restriction)
- \( c[p] \oplus (l(v)).P \) (selection towards role \( p \))
- \( c[p] &_{i \in I} \{l_i(x_i).P_i\} \) (branching from role \( p \) — with \( I \neq \emptyset \))
- \( \text{def } D \in Q \mid X(\bar{\bar{x}}) \) (process definition, process call)

- Declarations: \( D ::= X(\bar{\bar{x}}) = P \) (process declaration)
- Channels: \( c ::= x \mid s[p] \) (variable, channel with role \( p \))
- Values: \( v ::= c \mid \text{false} | \text{true} | 42 \ldots \) (channel, base value)

\( \text{fc}(P) \) is the set of free channels with roles in \( P \), and \( \text{fv}(P) \) is the set of free variables in \( P \).

A channel \( c \) can be either a variable or a channel with role \( s[p] \), i.e., a multiparty communication endpoint whose user impersonates role \( p \) in the session \( s \). Values \( v \) can be variables, or channels with roles, or base values. The inaction \( 0 \) represents a terminated process. The parallel composition \( P \mid Q \) represents two processes that can execute concurrently, and potentially communicate. The session restriction \( (\nu s)P \) declares a new session \( s \) with scope limited to process \( P \). Process \( c[p] \oplus (l(v)).P \) performs a selection (internal choice) towards role \( p \), using the channel \( c \): the labelled value \( l(v) \) is sent, and the execution continues as process \( P \). Dually, process \( c[p] &_{i \in I} \{l_i(x_i).P_i\} \) uses channels \( c \) to wait for a branching (external choice) from role \( p \): if the labelled value \( l_k(v) \) is received (for some \( k \in I \)), then the execution continues as \( P_k \) (with \( x_k \) holding value \( v \)). Note that for all \( i \in I \), variable \( x_i \) is bound with scope \( P_i \). In both branching and selection, the labels \( l_i \) (\( i \in I \)) are all different and their order is irrelevant. Process definition \( \text{def } D \in Q \) and process call \( X(\bar{\bar{x}}) \) model recursion, with \( D \) being a process declaration \( X(\bar{\bar{x}}) = P \): the call invokes \( X \) by expanding it into \( P \), and replacing its formal parameters with the actual ones. We postulate that process declarations are closed, i.e., in \( X(\bar{\bar{x}}) = P \), we have \( \text{fv}(P) \subseteq \bar{\bar{x}} \) and \( \text{fc}(P) = \emptyset \). Note that our syntax is simplified in the style of [19]: it does not have dedicated input/output prefixes, but they can be easily encoded using \& (with one branch) and \oplus.

**Example 2.2.** The following MPST π-calculus process implements the scenario in Figure 1:

\[
\begin{align*}
\text{def } \text{Loop}_b(x) &= x[a] \& \{\text{noviv}(y).x[c] \oplus (\nu s)[e](y)).\text{Loop}_b(x), \text{noviv}(z).x[c] \oplus (\nu s)[e](z)).\text{Loop}_b(x)\} \text{ in} \\
\text{def } \text{Client}_b(y) &= y[a] \& \text{play}(y).z[c] \oplus (\nu s)[e]("...").z[a] \& (\nu s)[e](y).\text{Loop}_b(z) \text{ in} \\
(\nu s_b, s_a, s_c) &\{Q \mid P_b \mid P_c\}
\end{align*}
\]

where: \( P_b = \text{Client}_b(s_b[p]) \) (for brevity, we omit the definitions of \( P_b \) and \( P_c \))

\[
Q = (\nu s)(s_a[q] \oplus (\nu s)[e](s[a]) \mid s_b[q] \oplus (\nu s)[e](s[b]) \mid s_c[q] \oplus (\nu s)[e](s[c]))
\]
A Linear Decomposition of Multiparty Sessions

In the 3rd line, $s_a, s_b, s_c$ are the sessions between the server process $Q$ and the clients $P_a, P_b, P_c$, which are composed in parallel with $|$. Each session has 2 roles: $q$ (server) and $p$ (client); e.g., $s_a$ is accessed by the server (through the channel with role $s_a[q]$) and by the client $P_a$ (through $s_a[p]$); similarly, $s_b$ (resp. $s_c$) is accessed by $P_b$ (resp. $P_c$) through $s_b[p]$ (resp. $s_c[p]$), while the server owns $s_a[q]$ (resp. $s_c[q]$). The body of the server process $Q$ defines a session $s$ (with 3 roles $a, b, c$) for playing the game. Note that the scope of $s$ does not include $P_a, P_b, P_c$: see Figure 2(a) for a schema of processes and sessions.

The server $Q$ uses the channel with role $s_a[q]$ (resp. $s_a[q], s_c[q]$) to send the message $\text{Play}^q$ (resp. $\text{Play}^q, \text{Play}^c$) carrying the channel with role $s[b]$ (resp. $s[a], s[c]$) to $p$. The result is a delegation of the channel to the client process $P_b$ (resp. $P_a, P_c$). This way, each client obtains a channel endpoint to interact in the game session $s$, interpreting a role among $a, b, c$.

The client $P_b$ is implemented by invoking $\text{Client}_b(s_a[p])$ (defined in the 2nd line). Here, $y[q]\&^{\text{Play}^q}(z)$ (defined in the 1st line), which becomes $\text{Loop}_b(z)$ after delegation.

**Definition 2.3.** The operational semantics of multiparty session processes is:

(R-Comm) $s[p][q]\&_{i\in I}\{l_i(x_i).P_i\} | s[q]\oplus \{l_j(y_j).Q\} \rightarrow P_i\{x_i/j\} | Q$ (if $j \in I$ and $\text{fv}(v) = \emptyset$)

(R-Call) $\text{def } X(\bar{x}) = P \in (X(\bar{v}) \mid Q) \rightarrow \text{def } X(\bar{z}) = P \in (P[\bar{v}/\bar{z}] \mid Q)$ (if $\bar{z} = x_1, \ldots, x_n, \bar{v} = v_1, \ldots, v_n$, $\text{fv}(\bar{v}) = \emptyset$)

(R-Par) $P \rightarrow Q$ implies $P | R \rightarrow Q | R$ (R-Res) $P \rightarrow Q$ implies $(\nu s)P \rightarrow (\nu s)Q$

(R-Def) $P \rightarrow Q$ implies $\text{def } D \text{ in } P \rightarrow \text{def } D \text{ in } Q$

(R-Struct) $P \equiv P'$ and $P \rightarrow Q$ and $Q' \equiv Q$ implies $P' \rightarrow Q'$ (with $\equiv$ standard — see [60])

Rule (R-Comm) models communication: it says that the parallel composition of a branching and a selection process, both operating on the same session $s$ respectively as roles $p$ and $q$ (i.e., via $s[p]$ and $s[q]$) and targeting each other (i.e., $s[p]$ is used to branch from $q$, and $s[q]$ is used to select towards $p$) reduces to the corresponding continuations, with a value substitution on the receiver side. (R-Call) says that a process call $X(\bar{v})$ in the scope of $\text{def } X(\bar{z}) = P \in \ldots$ reduces by expanding $X(\bar{v})$ into $P$, and replacing the formal parameters $\bar{z}$ with the actual ones $\bar{v}$. The remaining rules are standard: reduction can happen under parallel composition, restriction and process definition. By (R-Struct), reduction is closed under a structural congruence [60] stating, e.g., that $|$ is commutative and associative, and has $0$ as neutral element (i.e., $P | Q \equiv Q | P$, $P | (Q | R) \equiv (P | Q) | R$ and $P | 0 \equiv P$).

**Example 2.4.** The process in Example 2.2 reduces as (see also Figure 2(b), noting the scope of $s$):

$(\nu s_a, s_b, s_c)(Q | P_a | P_b | P_c) \rightarrow$ (by (R-Comm) between $Q$ and $P_a$ (R-Par), (R-Struct), (R-Res))

$(\nu s_a, s_c)(\nu s_b)(s_a[q][p]\oplus \text{Play}^q(s[a]) | s_c[q][p]\oplus \text{Play}^c(s[c])) | s[b][c]\oplus \text{Info}(\ldots) | P_a | P_c)$

**2.1 Multiparty Session Typing**

We now illustrate the typing system for the MPST $\pi$-calculus, and its properties. We adopt standard definitions from literature — except for some crucial (and duly noted) adaptations.

The goal of the MPST typing system is to ensure that processes interact on their channels according to given specifications, represented as session types. MPST foster a top-down approach: a global type $G$ describes a protocol involving various roles — e.g., the game with roles $a, b, c$ in Section 1; $G$ is projected into a set of (local) session types $S_a, S_b, S_c, \ldots$ (one per
role) that specify how each role is expected to use its channel endpoint; finally, session types are assigned to channels, and the processes using them are type-checked. Typing ensures that processes (1) never go wrong (i.e., use their channels type-safely), and (2) interact according to $G$, by respecting its projections — thus realizing a multiparty, deadlock-free session.

In the following, we provide a revised and streamlined presentation that clearly outlines the interplay between the global/local typing levels. For this reason, unlike most papers, we discuss local types first, and global types later, at the end of the section.

**Session Types: Local and Partial.** Session types describe the expected usage of a channel, as a communication protocol involving two or more roles. They allow to declare structured sequences of input/output actions, specifying who is the source/target role of interaction.

- **Definition 2.5 (Types and roles).** The syntax of (local) session types is:

  $$S ::= p \& \{i \in I \mid \exists l_i \in S_i\} (\text{branching from role } p \text{ — with } I \neq \emptyset)$$

  $$p \oplus \{i \in I \mid \exists l_i \in S_i\} (\text{selection towards role } p \text{ — with } I \neq \emptyset)$$

  $$\mu t.S \mid t \mid \text{end} (\text{recursive type, type variable, termination})$$

  $$B ::= \text{Bool} \mid \text{Int} \ldots (\text{base type}) \quad U ::= B \mid S (\text{closed}) (\text{payload type})$$

  We omit $\&/\oplus$ when $I$ is a singleton: $p.l_i(\text{Int}).S_i$ stands for $p \oplus i \in I \mid l_i(\text{Int}).S_i$.

  The set of roles in $S$, denoted as roles($S$), is defined as follows:

  $$\text{roles}(p \& i \in I \mid l_i(\text{Int}).S_i) \triangleq \text{roles}(p \& i \in I \mid l_i(\text{Int}).S_i) \triangleq \{p\} \cup \bigcup_{i \in I} \text{roles}(S_i)$$

  $$\text{roles}(\text{end}) \triangleq \emptyset \quad \text{roles}(t) \triangleq \emptyset \quad \text{roles}(\mu t.S) \triangleq \text{roles}(S)$$

  We will write $p \in S$ for $p \in \text{roles}(S)$, and $p \in S \setminus q$ for $p \in \text{roles}(S) \setminus \{q\}$.

  The **branching type** $p \& i \in I \mid l_i(\text{Int}).S_i$ describes a channel that can receive a label $l_i$ from role $p$ (for some $i \in I$, chosen by $p$), together with a payload of type $U_i$; then, the channel must be used as $S_i$. The **selection** $p \oplus i \in I \mid l_i(\text{Int}).S_i$, describes a channel that can choose a label $l_i$ (for any $i \in I$), and send it to $p$ together with a payload of type $U_i$; then, the channel must be used as $S_i$. The labels of branch/select types are all distinct and their order is irrelevant. The **recursive type** $\mu t.S$ and **type variable** $t$ model infinite behaviours. **end** is the type of a **terminated channel** (often omitted). **Base types** $B, B', \ldots$ can be types like $\text{Bool}, \text{Int}$, etc. **Payload types** $U, U', \ldots$ are either base types, or **closed session types**.

- **Example 2.6.** See the definition and description of session type $S_b$ in Section 1 (p. 3).

  To define session typing contexts later on, we also need **partial session types**.

- **Definition 2.7.** Partial session types, denoted by $H$, are:

  $$H ::= \& i \in I \mid l_i(\text{Int}).H_i \mid \oplus i \in I \mid l_i(\text{Int}).H_i (\text{branching, selection)} (\text{with } I \neq \emptyset, U_i \text{ closed})$$

  $$\mu t.H \mid t \mid \text{end} (\text{recursive type, type variable, termination})$$

  A partial session type $H$ is either a branching, a selection, a recursion, a type variable, or a terminated channel type. Unlike Definition 2.5, partial types have no role annotations: they are similar to **binary session types** (but the payloads $U_i$ can be multiparty) — and similarly, they endow a notion of **duality**: the outputs of a type match the inputs of its dual, and vice versa.

- **Definition 2.8.** $\overline{H}$ is the dual of $H$, defined as:

  $$\& i \in I \mid l_i(\text{Int}).\overline{H_i} \triangleq \& i \in I \mid l_i(\text{Int}).\overline{H_i}$$

  $$\oplus i \in I \mid l_i(\text{Int}).\overline{H_i} \triangleq \oplus i \in I \mid l_i(\text{Int}).\overline{H_i}$$

  $$\text{end} \triangleq \text{end} \quad t \triangleq t \quad \mu t.H \triangleq \mu t.\overline{H}$$
The dual of a selection type is a branching with dualised continuations, and *vice versa*; the payloads $U_i$ are the same. Duality is the identity on `end` and $t$, and homomorphic on $\mu_t.H$.

Multiparty session types can be *projected* onto a role $q$ (Definition 2.9 below): this yields a partial type that only describes the communications where $q$ is involved. This is technically necessary for typing rules, as we will see in Definition 2.11 later on.

**Definition 2.9.** $S \upharpoonright q$ is the partial projection of $S$ onto $q$:

\[
\begin{align*}
\text{end} \upharpoonright q & \triangleq \text{end} & t \upharpoonright q & \triangleq t & (\mu t.S) \upharpoonright q & \triangleq \begin{cases} 
\mu t.(S \upharpoonright q) & \text{if } S \upharpoonright q \neq q' \quad (\forall t') \\
\text{end} & \text{otherwise}
\end{cases} \\
(p \oplus_{i \in I} ?i(U_i).S_i) \upharpoonright q & \triangleq \begin{cases} 
\oplus_{i \in I} ?i(U_i).S_i \upharpoonright q & \text{if } p = p, \\
\Pi_{i \in I} (S_i \upharpoonright q) & \text{if } p \neq q
\end{cases} \\
(p \&_{i \in I} !i(U_i).S_i) \upharpoonright q & \triangleq \begin{cases} 
\&_{i \in I} !i(U_i).S_i \upharpoonright q & \text{if } q = p, \\
\Pi_{i \in I} (S_i \upharpoonright q) & \text{if } p \neq q
\end{cases}
\end{align*}
\]

where $\Pi$ is the *merge operator* for partial session types:

\[
\begin{align*}
\text{end} \cap \text{end} & \triangleq \text{end} & t \cap t & \triangleq t & \mu t.H \cap \mu t.H' & \triangleq \mu t.(H \cap H') & \&_{i \in I} ?i(U_i).H_i \cap \&_{i \in I} ?i(U_i).H_i' & \triangleq \&_{i \in I} ?i(U_i)(H_i \cap H_i') \\
(p \oplus_{i \in I} ?i(U_i).S_i) \cap (p \oplus_{i \in I} ?i(U_i).S_i) & \triangleq (\oplus_{i \in I} ?i(U_i).S_i) \cap (\oplus_{i \in I} ?i(U_i).S_i) \\
(p \&_{i \in I} !i(U_i).S_i) \cap (p \&_{i \in I} !i(U_i).S_i) & \triangleq (\&_{i \in I} !i(U_i).S_i) \cap (\&_{i \in I} !i(U_i).S_i)
\end{align*}
\]

The projection of `end` or a type variable $t$ onto any role is the identity. Projecting a recursive type $\mu t.S$ onto $q$, means projecting $S$ onto $q$, if $S \upharpoonright q$ is *not* some $q'$, for all possible recursive variables $t'$; otherwise, the projection is `end`. The projection of a selection $p \oplus_{i \in I} ?i(U_i).S_i$ (resp. branching $p \&_{i \in I} !i(U_i).S_i$) on role $p$, produces a partial selection type $\&_{i \in I} !i(U_i).S_i$ (resp. branching $\&_{i \in I} ?i(U_i).S_i$) with the continuations projected on $p$. Otherwise, if projecting on $q \neq p$, the select/branch is “skipped,” and the projection is the *merging of the continuations*, i.e., $\Pi_{i \in I} (S_i \upharpoonright q)$. The $\cap$ operator (introduced in [67, 18]) expands the set of session types whose partial projections are defined, which allows to type more processes (as we will see in Definition 2.11 and Example 2.14 later on). Crucially, $\cap$ can compose different *internal* choices, but *not* external choices (because this could break type safety).

**Subtyping.** The *subtyping relation* (Definition 2.10) says that a session type $S$ is “smaller” than $S'$ when $S$ is “less demanding” than $S'$ — i.e., when $S$ permits more internal choices, and imposes less external choices, than $S'$. When typing processes (Definition 2.12), a channel with a smaller type can be used whenever a channel with a larger type is required, according to Liskov’s Substitution Principle [45]. Subtyping is defined on both local and partial types.

**Definition 2.10 (Subtyping).** The *subtyping $\leq_S$ on multiparty session types* is the largest relation such that

(i) if $S \leq_S S'$, then $(\forall p \in (\text{roles}(S) \cup \text{roles}(S')))(S \upharpoonright \! p) \leq_{\! p} S' \upharpoonright \! p$, and

(ii) is closed backwards under coinductive rules at the top of Figure 3.

The *subtyping $\leq_p$ on partial session types* is coinductively defined by the rules at the bottom of Figure 3.

Definition 2.10 uses coinduction to support recursive types [56, Section 20 and Section 21]. Clause (i) links local and partial subtyping, and ensures that if two types are related, then their partial projections exist: this will be necessary later, for typing contexts (Definition 2.11). The gist of Definition 2.10 lies in clause (ii). Rules $(\text{s-Boca})/(\text{s-Sel})$ define subtyping on
branch/select types. Both rules are covariant in the continuation types, i.e., they require \( S_i \leq S_i' \). \((S\text{-Brch})\) is covariant also in the number of branches offered, whereas \((S\text{-Sel})\) is contravariant. \((S\text{-B})\) relates base types, if they are related by \( \leq B \). \((S\text{-End})\) relates terminated channel types. \((S\text{-μL})\) and \((S\text{-μR})\) are standard under coinduction: they say that a recursive session type \( \mu.T.S \) is related to \( S' \), iff its unfolding is related, too. The subtyping \( \leq P \) for partial types is similar, except for the lack of role annotations (thus resembling the binary session subtyping \([22]\)).

### Multiparty Session Typing System.

Before delving into the session typing rules (Definition 2.12), we need to formalise the notions of typing context and typing judgement, defined below.

**Definition 2.11.** A session typing context \( \Gamma \) is a partial mapping defined as:

\[
\Gamma ::= \emptyset \mid \Gamma, x: U \mid \Gamma, s[p]: S \text{ (with } p \not\in S) \tag{S-ContCont}
\]

We say that \( \Gamma \) is consistent iff for all \( s[p]: S_p, s[q]: S_q \in \Gamma \) with \( p \neq q \), we have \( S_p \upharpoonright q \leq P S_q \upharpoonright p \).

We say that \( \Gamma \) is complete if for all \( s[p]: S_p \in \Gamma, q \in S_p \) implies \( s[q] \in \text{dom}(\Gamma) \). We say that \( \Gamma \) is unrestricted, \( \text{un}(\Gamma) \), iff for all \( c \in \text{dom}(\Gamma) \), \( \Gamma(c) \) is either a base type or \( \text{end} \). The typing contexts composition \( \circ \) as neutral element:

\[
\begin{align*}
\Gamma_1, c: U \circ \Gamma_2, c': U' & \equiv (\Gamma_1 \circ \Gamma_2), c: U, c': U' \quad \text{(if } \text{dom}(\Gamma_2) \not= c \not= c' \not\in \text{dom}(\Gamma_1)) \\
\Gamma_1, x: B \circ \Gamma_2, x: B & \equiv (\Gamma_1 \circ \Gamma_2), x: B
\end{align*}
\]

A typing context can map a channel with role \( s[p] \) to a session type \( S \) (that cannot refer to \( p \) itself, ruling out “self-interactions”), but not to a base type. Variables can be mapped to either session or base types. The clause “\( \forall c: S \in \Gamma : S \not\in P p \)” is defined below.

**On Consistency.** In Definition 2.11, and in the rest of this work, we emphasise the importance of consistency of the context \( \Gamma \) for session typing: this condition is, in fact, **necessary to prove subject reduction**, and will be central for our encoding (Section 5 and Section 6).

As an example of non-consistent typing context, consider \( s[p]: \text{end}, s[q]: p?!(U).S \): we have \( \text{end} \upharpoonright q = \text{end} \not\leq P \text{?!(U).S} = (p?!(U).S) \upharpoonright p \).

Note that our consistency in Definition 2.11 is **weaker** than the one in previous papers (where it is sometimes called **coherency**): we use \( \leq P \), instead of (syntactic) type equality \( = \), to relate dual partial projections. The reason being: if we use \( = \), and adopt partial projections with type merging (Definition 2.9), subject reduction does **not** hold. Hence, by
Typing judgements

The judgement \( \Gamma \vdash \cdot \) is well typed if both \( \Theta \vdash \cdot \) and \( \Gamma \vdash \cdot \) are well typed under the composition of the corresponding typing contexts, as per Definition 2.11. By (T-Def), a process definition \( \text{def } X(\vec{x}) = P \) in \( Q \) is well typed if both \( P \) and \( Q \) are well typed in their typing contexts enriched with \( \vec{x} : \vec{U} \). For convenience, we annotate \( \vec{U} \) with types \( \vec{U} \), and \( \vec{U} \) with \( \vec{U} \) and \( \vec{U} \) with \( \vec{U} \). By (T-Call), process call \( X(v_1, \ldots, v_n) \) is well typed if the actual parameters \( v_1, \ldots, v_n \) have compatible types w.r.t. \( X \).

As mentioned above, we emphasise consistency by restricting typing judgements to consistent typing contexts — i.e., those allowing to prove subject reduction. The clause ”\( \forall c : S \vdash \cdot \) is defined” is unusual in MPST works, but arises naturally: by requiring the existence of partial projections, it rejects processes containing

(a) a channel with role \( s[p] : S \) that, for some \( q \in S \), cannot be (consistently) paired with \( s[q] \), or

(b) a variable \( x : S \) that, in a consistent and complete \( \Gamma \), cannot be substituted by any \( s[p] : S \).

Rejected processes cannot join any complete session (case (a)), or are never-executed “dead code” (case (b)).
Remark 2.13. Unlike most MPST papers (e.g., [19, 11]), our rule (T-Res) does not directly map a session \( s \) to a global type: this is explained in the next section, “Global Types”.

Example 2.14. Consider the session type \( S_b \) in Section 1 (p. 3), and the client process \( P_b = \text{Client}_b(s_b[p]) \) from Example 2.2. By Definition 2.12, the following typing judgement holds:

\[
\text{Client}_b:: q?\text{PlayB}(S_b), \text{Loop}_b:: \mu t. a \& \{ \text{Mov1ABC}(\text{int}), c!\text{Mov2BC}(\text{int}), t \} \cdot s_b[p]:: q?\text{PlayB}(S_b) \vdash \text{Client}_b(s_b[p])
\]

It says that the channel with role \( s_b[p] \) is used following type \( q?\text{PlayB}(S_b).\text{end} \) (with a delegation of a \( S_b \)-typed channel); the argument of \( \text{Client}_b \) has the same type; the argument of \( \text{Loop}_b \) is used following the game loop. This example cannot be typed without merging \( \sqcap \) (Definition 2.9): its derivation requires to compute

\[
S_b[c = !\text{InfoBC}(\text{string}), \mu t. (\text{Mov1BC}(\text{int}), t \sqcap \text{Mov2BC}(\text{bool})), t] = !\text{InfoBC}(\text{string}), \mu t. (\text{Mov1BC}(\text{int}), t \sqcap \text{Mov2BC}(\text{bool})), t)
\]

which is undefined without merging.

The typing rules in Figure 4 satisfy a subject reduction property (Theorem 2.16) based on typing context reductions. Reduction relations for typing contexts are common in typed process calculi, and reflect the communications required by the types in \( \Gamma \).

Definition 2.15 (Typing context reduction). The reduction \( \Gamma \rightarrow \Gamma' \) is:

\[
\begin{align*}
&\text{G} \vdash s[p]: S_p, s[q]: S_q \rightarrow s[p]: S_p, s[q]: S_q' &\text{if} &\begin{cases} \text{unf}(S_p) = q \sqcup_i j \text{U}(U_j), S_j & k \in I \\ \text{unf}(S_q) = p \sqcup_i j \text{U}(U_j), S_j' & k \in I & U_k \leq U_k' \\
\end{cases} \\
&\Gamma, c: U \rightarrow \Gamma', c: U' &\text{if} &\Gamma \rightarrow \Gamma' \text{ and } U \leq U' 
\end{align*}
\]

Our Definition 2.15 is a bit less straightforward than the ones in literature: it accommodates subtyping (hence, uses \( \leq S \)) and our iso-recursive type equality (hence, unfolds types explicitly).

Theorem 2.16 (Subject reduction). If \( \Theta \cdot \Gamma \vdash P \text{ and } P \rightarrow P' \), then \( \exists \Gamma': \Gamma' \rightarrow \Gamma' \text{ and } \Theta \cdot \Gamma' \vdash P' \).

Global Types. We conclude this section with global types, mentioned in Section 2.1 and Remark 2.13.

Definition 2.17. The syntax of global types, ranged over by \( G \), is:

\[
G ::= p \rightarrow q; \{ l_i(U_i).G_i \}_{i \in I} \quad (\text{interaction — with } U_i \text{ closed}) \\
\mu t.G | t | \text{end} \quad (\text{recursive type, type variable, termination})
\]

Type \( p \rightarrow q; \{ l_i(U_i).G_i \}_{i \in I} \) states that role \( p \) sends to role \( q \) one of the (pairwise distinct) labels \( l_i \) for \( i \in I \), together with a payload \( U_i \) (Definition 2.5). If the chosen label is \( l_j \), then the interaction proceeds as \( G_j \). Type \( \mu t.G \) and type variable \( t \) model recursion. Type \text{end} states the termination of a protocol. We omit the braces \( \{ \ldots \} \) from interactions when \( I \) is a singleton: e.g., \( a \rightarrow b:: l_i(U_i).G_i \) stands for \( a \rightarrow b:: \{ l_i(U_i).G_i \}_{i \in \{1\}} \).

Example 2.18. The following global type formalises the \text{Game} described in Section 1 and Figure 1:

\[
G_{\text{Game}} = b \rightarrow c:: \text{InfoBC}(\text{string}), c \rightarrow a:: \text{InfoCA}(\text{string}), a \rightarrow b:: \text{InfoBA}(\text{string})
\]

\[
\mu t.a \rightarrow b:: \begin{align*}
&\{ \text{Mov1BC}(\text{int}), b \rightarrow c:: \text{Mov1BC}(\text{int}), c \rightarrow a:: \{ \text{Mov1CA}(\text{int}), t \}, \text{Mov2CA}(\text{bool}), t \} \\
&\text{Mov2BA}(\text{Bool}), b \rightarrow c:: \text{Mov2BC}(\text{bool}), c \rightarrow a:: \{ \text{Mov1CA}(\text{int}), t \}, \text{Mov2CA}(\text{bool}), t \}
\end{align*}
\]
In MPST theory, a global type $G$ with roles $p_i$ ($i \in I$) is used to project\(^2\) a set of session types $S_i$ (one per role). E.g., projecting $G_{\text{Game}}$ in Example 2.18 onto $b$ yields the session type $S_b$ (p. 3). When all such projections $S_i$ are defined, and all partial projections of each $S_i$ are defined (as per Definition 2.9), then we can define the projected typing context of $G$:

$$
\Gamma_G = \{s[p_i] : S_i\}_{i \in I}
$$

where $\forall i \in I : S_i$ is the projection of $G$ onto $p_i$.

and $\Gamma_G$ can be shown to be:

(a) consistent and complete, i.e., can be used to type the session $s$ by rule (T-Res) (Figure 4), and

(b) deadlock-free, i.e.: $\Gamma_G \not\vdash^{s[i]} \Gamma_G' \not\vdash$ implies $\forall i \in I : \Gamma_G'(s[p_i]) = \text{end}$.

Similarly, it can be shown that $\Gamma_G$ reduces as prescribed by $G$.

Now, from observation (a) above, we can easily define a “strict” version of rule (T-Res) (Figure 4) in the style of [19, 11], where

1. the clause ‘$\Gamma'$ complete” is replaced with “$\Gamma'$ is the projected typing context of some $G$”, and

2. in the conclusion, the annotation $(\nu s : \Gamma')$ is replaced with $(\nu s : G)$.

Further, observation (b) allows to prove Theorem 2.19 below, as shown e.g. in [5]: a typed ensemble of processes interacting on a single $G$-typed session is deadlock-free (note: with our rules in Figure 4, the annotation $(\nu s : G)$ would be $(\nu s : \Gamma_G)$).

**Theorem 2.19** (Deadlock freedom). Let $\cdot \otimes \cdot \vdash P$, where $P \equiv (\nu s : G)_{i \in I} P_i$ and each $P_i$ only interacts on $s[p_i]$. Then, $P$ is deadlock-free: i.e., $P \not\vdash^{s[i]} P' \not\vdash$ implies $P' \equiv \emptyset$.

Note that the properties above emerge by placing suitable session types $S_i$ in the premises of (T-Res) — but our streamlined typing rules in Figure 4 do not require it, nor mention $G$. The main property of such rules is ensuring type safety (Theorem 2.16). We will exploit this insight (obtained by our separation of global/local typing) in our encoding (Section 5), preserving semantics and types (and thus, Theorem 2.19) without explicit references to global types.

### 3 Linear $\pi$-Calculus

The $\pi$-calculus is the canonical model for communication and concurrency based on message-passing and channel mobility. It was developed in the late 1980’s, with the first publication in 1992 [47], followed by various proposals for types and type systems. In this section we summarise the theory of the $\pi$-calculus with linear types [37], adopting a standard formulation and well-known results from [59]. We will present new $\pi$-calculus-related results in Section 4.

**Definition 3.1.** The syntax of $\pi$-calculus processes and values is:

\[
\begin{align*}
P, Q & ::= 0 \mid P \mid Q \mid (\nu x)P \quad \text{(inaction, parallel composition, restriction)} \\
& \quad \mid \tau(v), P \mid x(y).P \quad \text{(process replication, output, input)} \\
& \quad \mid \text{case v of } \{l_i(x_i) : P_i\}_{i \in I} \quad \text{(variant destruct)} \\
& \quad \mid \text{with } \{l_i : x_i\}_{i \in I} = \nu \text{ do } P \quad \text{(labelled tuple destruct)} \\
& \quad \mid u, v \ ::= x, y, w, z \mid l(v) \mid [l_i : v_i]_{i \in I} \quad \text{(name, variant value, labelled tuple value)} \\
& \quad \mid \text{false} \mid \text{true} \mid 42 \mid \ldots \quad \text{(base value)}
\end{align*}
\]

In $\pi$-calculus, names $x, y, \ldots$ can be intuitively seen as variables (i.e., they can be substituted with values), and as communication channels (i.e., they can be used for input/output). Values can be names, base values like $\text{false}$ or $42$, variant values $l(v)$ and labelled tuples

\(^2\) We use a standard projection with merging [67, 18]: for its definition (not crucial here), see [60].
The inaction \( 0 \) and the parallel composition \( P \parallel Q \) are similar to Definition 2.1. The restriction \( (\nu x)P \) creates a new name \( x \) and binds it with scope \( P \). The replicated process \( *P \) represents infinite replicas of \( P \), composed in parallel. The output \( \pi(v).P \) uses the name \( x \) to send a value \( v \), and proceeds as \( P \); the input \( x(y).P \) uses \( x \) to receive a value that will substitute \( y \) in the continuation \( P \). Process case \( v \) of \( \{l_i(x_i) \triangleright P_i\}_{i \in I} \) pattern matches a variant value \( v \), and if it has label \( l_i \), substitutes \( x_i \) and continues as \( P_i \). Process with \( [l_i : x_i]_{i \in I} \triangleright v \) do \( P \) destructs a labelled tuple \( v \), substituting each \( x_i \) in \( P \).

For brevity, we will often write “record” instead of “labelled tuple”.

**Definition 3.2.** The \( \pi \)-calculus operational semantics is the relation \( \rightarrow \) defined as:

\[
\begin{align*}
& (R_{\pi\text{-Com}}) \quad \pi(v).P \parallel x(y).Q \rightarrow P \parallel Q[\{\nu x\}] \\
& (R_{\pi\text{-Case}}) \quad \text{case}_{l_j(v)} \{l_i(x_i) \triangleright P_i\}_{i \in I} \rightarrow P_{j[\nu s_j]} \quad (j \in I) \\
& (R_{\pi\text{-Wrrn}}) \quad \text{with} \{l_i : x_i\}_{i \in I} = [l_i : v_i\]_{i \in I} do P \rightarrow P[\{\nu v_i\}]_{i \in I} \\
& (R_{\pi\text{-Rss}}) \quad P \rightarrow Q \text{ implies } (\nu x)P \rightarrow (\nu x)Q \\
& (R_{\pi\text{-Par}}) \quad P \rightarrow Q \text{ implies } P[R \rightarrow Q | R] \\
& (R_{\pi\text{-Struct}}) \quad P \equiv P' \land P \rightarrow Q \land Q' \equiv Q \text{ implies } P' \rightarrow Q'
\end{align*}
\]

Rule \((R_{\pi\text{-Com}})\) models communication between output and input on a name \( x \): it reduces to the corresponding continuations, with a value substitution on the receiver process. \((R_{\pi\text{-Case}})\) says that case applied on a variant value \( l_j(v) \) reduces to \( P_j \), with \( v \) in place of \( x_j \) — provided that \( l_j \) is one of the supported cases (i.e., \( l_j = l_i \) for some \( i \in I \)). Rule \((R_{\pi\text{-Wrrn}})\) deconstructs a labelled tuple \([l_i : v_i]_{i \in I} \): it says that with reduces to its continuation \( P \) with \( v_i \) in place of each \( x_i \), for all \( i \in I \). By \((R_{\pi\text{-Rss}})\) and \((R_{\pi\text{-Par}})\), reductions can happen under restriction and parallel composition, respectively. By \((R_{\pi\text{-Struct}})\), reduction is closed under the structural congruence \( \equiv \), whose definition is standard (see [59, Table 1.1] and [60]).

**\( \pi \)-Calculus Typing.** We now summarise the \( \pi \)-calculus types, subtyping, and typing rules.

**Definition 3.3 (\( \pi \)-types).** The syntax of a \( \pi \)-calculus type \( T \) is given by:

\[
T ::= \text{Li}(T) \mid \text{Lo}(T) \mid \text{L}^\ast(T) \mid (l_i.T_i)_{i \in I} \mid [l_i : T_i]_{i \in I} \mid \mu T.T \mid \bullet \mid \text{Bool} \mid \text{Int} \mid \ldots
\]

Linear types \( \text{Li}(T) \), \( \text{Lo}(T) \) denote, respectively, names used exactly once to input/output a value of type \( T \). \( \text{L}^\ast(T) \) denotes a name used once for sending, and once for receiving, a message of type \( T \). \( (l_i.T_i)_{i \in I} \) denotes an unrestricted connection, i.e., a name that can be used both for input/output any number of times. \( \bullet \) is assigned to names that cannot be used for input/output. \( (l_i.T_i)_{i \in I} \) is a labelled disjoint union of types, while \([l_i : T_i]_{i \in I} \) that we will often call “record” is a labelled product type; for both, labels \( l_i \) are all distinct, and their order is irrelevant. As syntactic sugar, we write \((T_i)_{i \in 1..n}\) for a record with integer labels \([i : T_i]_{i \in 1..n}\). Recursive types and variables, and base types like \text{Bool}, are standard.

The predicate \( \text{lin}(T) \) (Definition 3.4 below) holds iff \( T \) has some linear input/output component.

**Definition 3.4 (Linear/unrestricted types).** The predicate \( \text{lin} \) is inductively defined as:

\[
\text{lin}(\text{Li}(T)) \quad \text{lin}(\text{Lo}(T)) \quad 3j \in I : \text{lin}(T_j) \quad \exists j \in I : \text{lin}(T_j) \quad \text{lin}(\mu T.T) \quad \text{lin}((l_i.T_i)_{i \in I}) \quad \text{lin}([l_i : T_i]_{i \in I})
\]

We write \( \text{un}(T) \) iff \( \neg \text{lin}(T) \) (i.e., \( T \) is unrestricted if is not linear).
Figure 5 Typing rules for the linear \( \pi \)-calculus.

Definition 3.5. Subtyping \( \leq_{\pi} \) for \( \pi \)-types is coinductively defined as:

\[
\begin{align*}
B & \ll B' & (S-LB) \\
B & \leq_{\pi} B' & (S-LEND) \\
\forall i \in I \quad T_i & \leq_{\pi} T_i' & (S-VARIANT) \\
\{l_i : T_i\}_{i \in I} & \leq_{\pi} \{l_i : T_i\}_{i \in I} & (S-LTUPLE) \\
L_i(T) & \leq_{\pi} L_i(T') & (S-L) \\
L(T) & \leq_{\pi} L(T') & (S-Lo) \\
T & \leq_{\pi} T' & (S-Lp) \\
T(\mu t/v) & \leq_{\pi} T' & (S-Lp) \\
\mu T & \leq_{\pi} T & (S-LpL)
\end{align*}
\]

By rule \( (S-LB) \), \( \leq_{\pi} \) includes basic subtyping \( \leq_B \). \( (S-LEND) \) relates types without I/O capabilities. By \( (S-L) \) (resp. \( (S-Lo) \)), linear input (resp. output) subtyping is covariant (resp. contravariant) in the carried type. By \( (S-VARIANT) \), subtyping for variant types is covariant in both carried types and number of components. By \( (S-LTUPLE) \), subtyping for labelled tuples, a.k.a. records, is covariant in the carried types. (Note: “full” record subtyping allows to add/remove entries [59, §7.3]; but here, “record” just means “labelled tuple”.) Rule \( (S-LpL) \) (and its symmetric, omitted) relates a recursive type \( \mu t.T \) to \( T' \) iff its unfolding is related to \( T' \).

Definition 3.6 (Typing context, type combination). The linear \( \pi \)-calculus typing context \( \Gamma \) is a partial mapping defined as:

\[
\Gamma := \emptyset \mid \Gamma, x : T
\]

We write \( \text{lin}(\Gamma) \) iff \( \exists x : T \in \Gamma : \text{lin}(T) \), and \( \text{un}(\Gamma) \) iff \( \neg \text{lin}(\Gamma) \). The type combinator \( \uplus \) is defined as follows (and undefined in other cases), and is extended to typing contexts as expected.

\[
\text{lin}(T) \uplus \text{Lo}(T) \triangleq \text{L}_2(T) \quad \text{Lo}(T) \uplus \text{lin}(T) \triangleq \text{L}_2(T) \quad T \uplus T \triangleq T \quad \text{if } \text{un}(T)
\]

\[
(\Gamma_1 \uplus \Gamma_2)(x) \triangleq \begin{cases} \\
\Gamma_1(x) \uplus \Gamma_2(x) & \text{if } x \in \text{dom}(\Gamma_1) \cap \text{dom}(\Gamma_2) \\
\Gamma_1(x) & \text{if } x \in \text{dom}(\Gamma_1) \setminus \text{dom}(\Gamma_2) 
\end{cases}
\]

Figure 5 shows the typing system for the linear \( \pi \)-calculus. Typing judgements have two forms: \( \Gamma \vdash v : T \) and \( \Gamma \vdash P \). \( (\text{TName}) \) says that a name has the type assumed in the typing context; \( (\text{Basic}) \) relates base values to their types; both rules require unrestricted typing contexts. By \( (\text{LTVar}) \), a variant value \( l(v) \) is of type \( \{l_i : T_i\}_{i \in I} \) if value \( v \) is of type \( T \). By \( (\text{LPar}) \), a record value \( [l_i : v_i]_{i \in I} \) is of type \( [l_i : T_i]_{i \in I} \) if for all \( i \in I \), \( v_i \) is of type \( T_i \). \( (\text{SCon}) \) is the substitution rule: if \( x \) has type \( T \) in \( \Gamma \), then it also has any supertype of \( T \). By \( (\text{Nu}) \), \( \emptyset \) is well typed in every unrestricted typing context. By \( (\text{Par}) \), the parallel composition of two processes is typed by combining the respective typing contexts. By \( (\text{Res}) \), the restriction process \( \mu x.P \) is well typed if \( P \) is typed by augmenting the context with \( x : \text{Li}(T) \), or \( x : \not T \). In the first case, by applying Definition 3.6 (\( \uplus \)), we have \( x : \text{Li}(T) = x : \text{Li}(T) \uplus \text{Lo}(T) \); this implies that \( P \) owns both capabilities of linear input/output
We introduce some definitions and results on typed $\bot$ with what is the dual of $\mu$ variables (Definition 3.3) with their $\mu$-calculus counterparts, denoted with $\bar{\mu}$. We allow recursive types such as $\mu \bar{\mu} \bar{\mu} \bar{\mu} \bar{\mu}$ (but not $\bar{\mu} \bar{\mu} \bar{\mu} \bar{\mu}$), and postulate that when unfolding, $\bar{\mu}$ is substituted by a “dual” type $\mu \bar{\mu} \bar{\mu} \bar{\mu} \bar{\mu}$, as formalised in Definition 4.1 below. Quite interestingly, our

$$\text{let } x = v \text{ in } P \triangleq (\nu z) (T(z) \cdot 0) z(x), P \quad (\text{where } z \notin \{x\} \cup \text{fn}(v) \cup \text{fn}(P))$$

\[(\text{T-\text{Let}}) \quad \text{let } x = v \text{ in } P \rightarrow P[v/x]
\]

\[(\text{T-\text{Narrow}}) \quad \Gamma, x : T \vdash P \quad T' \leq T \quad \Gamma, x : T' \vdash P
\]

\[(\text{T-\text{MSubst}}) \quad \forall i \in I \quad \Gamma, i, v_i : T_i, \Gamma, \{x_i : T_i\}_{i \in I} \vdash P
\]

\[\frac{\Gamma \vdash v : T \quad \Gamma_2, x : T \vdash P}{\Gamma \vdash \text{let } x = v \text{ in } P} \quad (\text{T-\text{Let}})
\]

\[\text{Figure 6 “Let” binder (definition, reduction, typing), and narrowing / substitution rules.}
\]

of $x$. By (T-\text{Res}), the restriction $(\nu x)P$ is typed if $P$ is typed and $x$ has no capabilities. By (T-\text{Ins}) (resp. (T-Out)), the input and output processes are typed if $x$ is a (possibly linear) name used in input (resp. output), and the carried types are compatible with the type of $y$ (resp. value $v$). The typing context used to type the input and output process is obtained by applying $\otimes$ on the premises. By (T-\text{Repl}), a replicated process $*P$ is typed in the same unrestricted context that types $P$. By (T-\text{Case}), case $v$ of $[I_i (x_i) : P_i]_{i \in I}$ is typed if the guard value $v$ has variant type, and every $P_i$ is typed assuming $x_i : T_i$, for all $i \in I$. By (T-\text{Wtn}), process with $[I_i (x_i) : P_i] = v \Rightarrow P$ is typed if $v$ is of record type and for all $i \in I$, each $v_i$ has the same type as $x_i$, i.e., $T_i$.

4 Some Typed $\pi$-Calculus Extensions and Results

We introduce some definitions and results on typed $\pi$-calculus: we will need them in Section 5 and Section 6, to state our encoding and its properties. As we target standard typed $\pi$-calculus (Section 3), all our extensions are conservative, so to preserve standard results (e.g., subject reduction).

“Let” binder, narrowing, substitution. Figure 6 shows several auxiliary definitions and typing rules. let $x = v \text{ in } P$ binds $x$ in $P$, and reduces by replacing $x$ with $v$ in $P$. It is a macro on other $\pi$-calculus constructs: hence, rules (T-\text{Let})/(T-\text{Narrow}) are based on the reduction/typing of its expansion (details in [60]). Rule (T-\text{MSubst}) derives from the narrowing lemma [59, 7.2.5]. (T-\text{MSubst}) represents zero or more applications of the substitution lemma [59, 8.1.4].

Duality and Recursive $\pi$-Types. The duality for linear $\pi$-types relates opposite but compatible input/output capabilities. Intuitively, the dual of a $\text{Li}(T)$ is $\text{Lo}(T)$ (and vice versa) [15]. Note that the carried type $T$ is the same: i.e., dual types can be combined with $\otimes$ (Definition 3.6), yielding $\text{Li}(T)$. However, defining duality for recursive $\pi$-types is not straightforward: what is the dual of $T = \mu.t.\text{Lo}(t)$? Is it maybe $T' = \mu.t.\text{Li}(t)$? Since $\otimes$ is not defined for $\mu$-types, we can check whether it is defined for the unfoldings of our hypothetical duals $T$ and $T'$. Unfortunately, we have $\text{unf}(T) = \text{Lo}(\mu.t.\text{Lo}(t))$ and $\text{unf}(T') = \text{Li}(\mu.t.\text{Li}(t))$: i.e., $\otimes$ is again undefined, so $T, T'$ cannot be considered duals. Solving this issue is crucial: in Section 5, we will need to encode recursive partial types, preserving their duality (Definition 2.8) in linear $\pi$-types.

What we want is a notion of duality that commutes with unfolding, so that if two recursive types are dual, and we unfold them, we get a dual pair $\text{Lo}(T)/\text{Li}(T)$ that can be combined with $\otimes$ (since they carry the same $T$). We address this issue by extending the $\pi$-calculus type variables (Definition 3.3) with their dualised counterpart, denoted with $\bar{\pi}$. We allow recursive types such as $\mu.\bar{t} \cdot \text{Lo}(t)$ (but not $\mu.\bar{t} \cdot \ldots$), and postulate that when unfolding, $\bar{\pi}$ is substituted by a “dual” type $\mu.\bar{\mu} \cdot \text{Lo}(t)$, as formalised in Definition 4.1 below. Quite interestingly, our
approach reminds of the “logical duality” for session types [43], but we study it in the context of \(\pi\)-calculus (we will further discuss this topic in Section 8).

**Definition 4.1.** \(\overline{T}\) is the dual of \(T\), and is defined as follows:

\[
\begin{align*}
\text{Li}(T) & \triangleq \Lo(T) & \Lo(T) & \triangleq \text{Li}(T) & \bullet \triangleq \bullet & (\overline{t}) \triangleq \overline{t} & (\overline{\overline{t}}) \triangleq t & \mu t.T & \triangleq \mu t.(\overline{\overline{t}}) \\
\end{align*}
\]

The substitution of \(T\) for a type variable \(t\) or \(\overline{\overline{t}}\) is:

\[
(t^{T/y}) \triangleq T & \quad (\overline{\overline{t}}^{T/y}) \triangleq \overline{T}
\]

The dual of a linear input type \(\Li(T)\) is a linear output type \(\Lo(T)\), and vice versa, with the payload type \(T\) unchanged, as expected. The dual of a terminated channel type \(\bullet\) is itself. The dual of a type variable \(t\) is \(\overline{\overline{t}}\), and the dual of a dualised type variable \(\overline{\overline{t}}\) is \(t\), implying that duality on linear \(\pi\)-types is convolutive. The dual of \(\mu t.T\) is \(\mu t.(\overline{\overline{t}})\), where type \(T\) is dualised to \(\overline{T}\), and every occurrence of \(t\) is replaced by its dual \(\overline{\overline{t}}\) by Definition 4.1. Now, the desired commutativity between duality and unfolding holds, as per Lemma 4.2 below.

**Lemma 4.2.** \(\text{unf}(T) = \mu t.(\overline{\overline{t}})\).

**Example 4.3.** Let \(T = \mu t.\Li((t, \overline{\overline{t}}))\). Then:

\[
\begin{align*}
\text{unf}(T) &= \Li\left(\left(\mu t.\Li((t, \overline{\overline{t}})), \mu t.\Lo((\overline{\overline{t}}, t))\right)\right) = \Li\left(\left(\mu t.\Li((t, \overline{\overline{t}})), \mu t.\Lo((\overline{\overline{t}}, t))\right)\right) = \text{unf}(T)
\end{align*}
\]

By adding dualised type variables in Definition 3.3, we naturally extend the definition of \(\text{fv}(T)\) (with \(\mu \ldots\) binding both \(t\) and \(\overline{\overline{t}}\), the subtyping relation \(\leq_{\pi}\) in Definition 3.5 (by letting rules (s-LmL) and (s-LmR) use the substitution in Definition 4.1) and ultimately the typing system in Definition 3.6. Using these extensions, we will obtain a rather simple encoding of recursive session types (Definition 5.1), and solve a subtle issue involving duality, recursion and continuations (Example 5.3).

The reader might be puzzled about the impact of dualised variables in the \(\pi\)-calculus theory. We show that dualised variables do not increase the expressiveness of linear \(\pi\)-types, and do not unsafely enlarge subtyping \(\leq_{\pi}\): this is proved in Lemma 4.4, that allows to erase dualised variables from recursive \(\pi\)-types. It uses

1. a substitution that only replaces dualised variables, i.e.: \(\overline{\overline{t'}}/\overline{\overline{t}} = t'\); and
2. the equivalence \(=_{\pi}\) defined as: \(\leq_{\pi} \cap \leq_{\pi}^{-1}\).

**Lemma 4.4 (Erasure of \(\overline{\overline{t}}\)).** \(\mu t.T =_{\pi} \mu t.\overline{\overline{t'}}(\overline{\overline{t'}})/\overline{\overline{t}}, \text{ for all } t' \notin \text{fv}(T)\).

**Example 4.5 (Application of erasure).** Take \(T\) from Example 4.3. By Lemma 4.4, we have:

\[
T =_{\pi} \mu t.\Li\left(\left(t, \mu t'.\Li((t, \overline{\overline{t'}}))^{t'/t}\right)\right) = \mu t.\Li((t, \mu t'.\Lo((t, t')))).
\]

Since \(T =_{\pi} T'\) implies \(T \leq_{\pi} T'\) and \(T' \leq_{\pi} T\), Lemma 4.4 says that any \(\mu t.T\) is equivalent to a \(\mu\)-type without occurrences of \(\overline{\overline{t'}}\); i.e., any typing relation with instances of \(\overline{\overline{t}}\) corresponds to a \(\overline{\overline{t}}\)-free one. As a consequence, any typing derivation using \(\overline{\overline{t}}\) can be turned into a \(\overline{\overline{t}}\)-free one. Summing up: adding dualised variables preserves the standard results of typed \(\pi\)-calculus.

**Type Combinator \(\square\).** Definition 4.6 introduces a type combinator that is a “relaxed” version of \(\Box\) (Definition 3.6) extended with subtyping. We will use it to encode MPST typing contexts (Definition 5.6).

**Definition 4.6.** The \(\pi\)-calculus type combinator \(\square\) is defined on \(\pi\)-types as follows (and undefined in other cases), and naturally extended to typing contexts:

\[
\begin{align*}
\Lo(T) \cap \Li(T') & \triangleq \Li(T) \cup \Lo(T) & \Li(T') \cap \Lo(T) & \triangleq \Li(T') \cup \Lo(T) & \text{if } T \leq_{\pi} T' & T \cap T' =_{\pi} T & \text{if } \text{unf}(T)
\end{align*}
\]

\[
(\Gamma_1 \cap \Gamma_2)(x) \triangleq \begin{cases} 
\Gamma_1(x) \cap \Gamma_2(x) & \text{if } x \in \text{dom}(\Gamma_1) \cap \text{dom}(\Gamma_2) \\
\Gamma_1(x) & \text{if } x \in \text{dom}(\Gamma_1) \setminus \text{dom}(\Gamma_2)
\end{cases}
\]
The difference between υ and ρ is that the former combines linear inputs/outputs with the same carried type, while ρ is more relaxed: it allows a carried type to be subtype of the other — more exactly, the type carried by the output side can be smaller than the type carried by the input side. This is shown in Lemma 4.7 and Example 4.8 below.

Lemma 4.7. If \(T = T_1 \cap T_2\), and \(T_1 \cup T_2 = T\), then either
(a) \(T_1 \subseteq \pi T_1\) and \(T_2 \subseteq \pi T_2\), or
(b) \(T_1 \subseteq \pi T_1\) and \(T_2 \subseteq \pi T_2\).

Lemma 4.7 says that \(T_1 \cap T_2\) (when defined) is a type that, when split using υ, yields linear I/O types that are subtypes of the originating \(T_1, T_2\). Intuitively, it means that \(\rho\) can be soundly used to simplify typing derivations: if used to type some name \(x\), it will yield (when defined) a type that can also be obtained by suitably using υ and \((x_{\pi-.SUB})\) (Figure 5).

Example 4.8. Let \(T_1 = \text{Li}(\text{Real})\), \(T_2 = \text{Lo}(\text{Int})\), and \(T = T_1 \cap T_2\). We have \(T = \text{L}(\text{Int})\); if we let \(T_1 \cup T_2 = T\), then we get either (a) \(T_1 = \text{Li}(\text{Int})\) \(\subseteq \pi T_1\) and \(T_2 = \text{Lo}(\text{Int})\) \(\subseteq \pi T_2\), or (b) \(T_1 = \text{Lo}(\text{Int})\) \(\subseteq \pi T_2\) and \(T_2 = \text{Li}(\text{Int})\) \(\subseteq \pi T_1\).

5 Encoding Multiparty Session-π into Linear π-Calculus

We now present our encoding of MPST π-calculus into linear π-calculus. It consists of an encoding of types and an encoding of processes: combined, they preserve the safety properties of MPST communications, both w.r.t. typing and process behaviour.

Encoding of Types. Our goal is to decompose MPST channel endpoints into point-to-point π-calculus channels. This leads to the main intuition behind our approach: encode MPST channel endpoints as labelled tuples, whose labels are roles, and whose values are names (for communication). The idea is that if a multiparty channel of type \(S\) allows to talk with role \(p\), then the corresponding π-calculus record should have a label \(p\), mapping to a name that can send/receive messages to/from the process that plays the role \(p\). This suggests the type of an encoded MPST channel endpoint: it should be a π-calculus record — and since each name appearing in such record is used to communicate, it should have an input/output type.

Definition 5.1. The encoding of session type \(S\) into linear π-types is:

\[ [S] \triangleq [p : [S \mid p]]_{p \in S} \]

where the encoding of the partial projections \([S \mid p]\) is:

\[ \llbracket \psi \downarrow p \RRbracket. H_i \triangleq \text{Lo}\{\llbracket \llbracket \psi |- (\llbracket U_i \RRbracket, \llbracket H_i \RRbracket) \rrbracket \rrbracket\}_{i \in I} \]
\[ \llbracket \kappa \downarrow p \RRbracket. H_i \triangleq \text{Li}\{\llbracket \llbracket \kappa |- (\llbracket U_i \RRbracket, \llbracket H_i \RRbracket) \rrbracket \rrbracket\}_{i \in I} \]
\[ [B] \triangleq B \quad [\text{end}] \triangleq \bullet \]
\[ [t] \triangleq t \quad [\text{id}. H] \triangleq \text{id}. [H] \]

The encoding of a session type \(S\), namely \([S]\), is a record that maps each role \(p \in S\) to the encoding of the partial projection \([S \mid p]\). The latter adopts the basic idea of the encoding of binary, non-recursive session types [36, 15]: it is the identity on a base type \(B\), while a terminated channel type \(\text{end}\) becomes \(\bullet\), with no capabilities. Selection \(\psi \downarrow p\) \((\llbracket U_i \RRbracket, \llbracket H_i \RRbracket)\) is encoded as linear output and input types, respectively, adopting a continuation-passing style (CPS). In both cases, the carried types are variants: \(\llbracket \psi |- (\llbracket U_i \RRbracket, \llbracket H_i \RRbracket) \rrbracket_{i \in I}\) for select and \(\llbracket \kappa |- (\llbracket U_i \RRbracket, \llbracket H_i \RRbracket) \rrbracket_{i \in I}\) for branch, with the same labels as the originating partial projections. Such variants carry tuples \((\llbracket U_i \RRbracket, \llbracket H_i \RRbracket)\) and \((\llbracket U_i \RRbracket, \llbracket H_i \RRbracket):\) the first element is the encoded payload type, and the second (i.e., the encoding of \(H_i\)) is the type of a continuation name: it is sent together with the encoded payload, and will be used to send/receive the next message (unless \(H_i\) is \text{end}). Note that selection sends the dual of \([H_i]\): this is because the sender must keep interacting according to \([H_i]\), while the
recipient must operate dually (cf. Definition 4.1). E.g., if \( [H_i] \) requires to send a message, the recipient of \( [T_i] \) must receive it. The encodings of type variables and recursive types are homomorphic.

Note that by encoding session types as labelled tuples, we untangle the order of the interactions among different roles. We will recover this order later, when encoding processes.

**Example 5.2.** Consider the session type \( S \triangleq p | l_1(\text{Int}), q | ? l_2(S').\text{end} \), where \( S' \triangleq r | l_3(\text{Boolean}).q ? l_4(\text{String}).\text{end} \). By Definition 5.1, the encoding of \( S \) is:

\[
[S] = \{ p: [S | p], q: [S | q] \} = \{ p: [[l_1(\text{Int})], q: [? l_2(S')]] \}
\]

\[
= \{ p: \text{Lo}(l_1((\text{Int})), q: \text{Li}(l_2((r: \text{Lo}(l_3(\text{Boolean})), q: \text{Li}(l_4((\text{String})), \text{end})))))) \}
\]

**Recursion, Continuations and Duality.** We now point out a subtle (but crucial) difference between Definition 5.1 and the encoding of binary, non-recursive session types in [15]. When encoding partial selections, our continuation type is the dual of the encoding of \( H_i \), i.e., \( [H_i] \); in [15], instead, it is the encoding of the dual of \( H_i \), i.e., \( \overline{[H_i]} \). This difference is irrelevant for non-recursive types (Example 5.2); but for recursive types, using \( [H_i] \) would yield the wrong continuations. Using \( \overline{[H_i]} \), instead, gives the expected result, by generating dualised recursion variables (cf. Definition 4.1). We explain it in Example 5.3 below.

**Example 5.3.** Let \( H = \mu t. ! t!(\text{Boolean}).t \). By Definition 5.1, we have:

\[
[H] = [\mu t. ! t!(\text{Boolean}).t] = \mu t. \text{Lo}(l_{\text{Boolean}}, t) = \mu t. \text{Lo}(l_{\text{Boolean}}, t)
\]

Let us now unfold the encoding of \( H \). By Definition 4.1, we have:

\[
\text{unf}([H]) = \text{unf}(\mu t. \text{Lo}(l_{\text{Boolean}}, t)) = \text{Lo}(l_{\text{Boolean}}, \mu t. \text{Li}(l_{\text{Boolean}}, t))
\]

This is what we want: since \( H \) requires a recursive output of \( \text{Boolean} \), its encoding should output a \( \text{Boolean} \) together with a recursive input name as continuation. Hence, the recipient will receive the first \( \text{Boolean} \) together with a continuation name, whose type mandates to recursively input more \( \text{Boolean} \). If encoding continuations as in [15], instead, we would have:

\[
[H] = \mu t. \text{Lo}(l_{\text{Boolean}}, t) = \mu t. \text{Lo}(l_{\text{Boolean}}, t) (t \text{ is not dualised})
\]

\[
\text{unf}([H]) = \text{Lo}(l_{\text{Boolean}}, \mu t. \text{Lo}(l_{\text{Boolean}}, t))
\]

which is wrong: the recipient is required to recursively output \( \text{Boolean} \). This wrong encoding would also prevent us from obtaining Theorem 6.1 later on.

**Encoding of Typing Contexts.** In order to preserve type safety, we want to encode a session judgement (Figure 4) into a \( \pi \)-calculus typing judgement (Figure 5). For this reason, we now use the encoding of session types (Definition 5.1) to formalise the encoding of session typing contexts.

**Definition 5.4.** The encoding of a session typing context is:

\[
\begin{align*}
\Theta, \Gamma &\triangleq \Theta, \Gamma \\
\Theta, X: \bar{U} &\triangleq (\Theta, [X: \bar{U}]) \\
\Gamma, c: U &\triangleq [c: U] \\
\Theta, c: U &\triangleq [c: U] \\
\Theta[\tau_p] &\triangleq \approx_p[\tau_p] \\
\hat{\Gamma}_1 \circ \hat{\Gamma}_2 &\triangleq [\hat{\Gamma}_1 \circ \hat{\Gamma}_2] \\
\hat{\Gamma}_1 \cup \hat{\Gamma}_2 &\triangleq [\hat{\Gamma}_1 \cup \hat{\Gamma}_2] \\
\hat{\Gamma}_1, \cdots, \hat{\Gamma}_n &\triangleq [\hat{\Gamma}_1, \cdots, \hat{\Gamma}_n] \\
\end{align*}
\]

When encoding typing contexts, variables \((x)\) keep their name, while process variables \((\approx_p)\) and channels with roles \((\tau_p)\) are turned into distinguished names with a subscript: e.g., \( X \) becomes \( z_X \). The composition \( \hat{\Gamma}_1 \circ \hat{\Gamma}_2 \) (Definition 2.11) is encoded using \( \cup \) (Definition 3.6); such an operation is always defined, since the domains of \( \hat{\Gamma}_1, \hat{\Gamma}_2 \) can only overlap on basic types.

Note that encoded process variables have an unrestricted connection type, carrying an \( n \)-tuple of encoded argument types; encoded sessions, instead, are linearly-typed, similarly
Encoding Typing Judgements: Overview. With these definitions at hand, we can now have a first look at the encoding of session typing judgements in Figure 7 (but we postpone the formal statement to Definition 5.7 later on, as it requires some more technical developments).

Terminated processes are encoded homomorphically. Parallel composition is also encoded homomorphically — i.e., our encoding preserves the choreographic distribution of the originating processes. Note that \([P]_\Theta \cdot \Gamma_1\) and \([Q]_\Theta \cdot \Gamma_2\) are the encoded processes yielded respectively by \([\Theta \cdot \Gamma_1 \vdash P]\) and \([\Theta \cdot \Gamma_2 \vdash Q]\): they exist because such typing judgements hold, by inversion of (T-Par) (Figure 4). Similar uses of sub-processes encoded w.r.t. their typing occur in the other cases. Process declaration \(\text{def} X(\bar{v} : U) = P \text{ in } Q\) is encoded as a replicated \(\pi\)-calculus process that inputs a value \(z\) on a name \(X = z_X\) (matching Definition 5.4), deconstructs it into \(x_1, \ldots, x_n\) (using with, and hence assuming that \(z\) is an \(n\)-tuple), and then continues as the encoding of the body \(P\); meanwhile, the encoding of \(Q\) runs in parallel, enclosed by a delimitation on \(z_X\) (that matches the scope of the original declaration). Correspondingly, a process call \(X(\bar{v})\) is encoded as a process that sends the encoded values \([\bar{v}]\) on \(z_X\) and ends (in MPST \(\pi\)-calculus, process calls are in tail position).

Selection on \(c[p]\) is encoded using information from the session typing context: the fact that \(c\) has type \(S = p \oplus \uplus (U).S'\) — i.e., \([S]\) is a record type with one entry \(q : z_q\) for each \(q \in S\). Therefore, the encoding first deconstructs \([c]\) (using with), and then uses the (linear) name in its \(p\)-entry to output on \(z_p\). Before performing the output, however, a new name \(z\) is created: it is the continuation of the interaction with \(p\). Then, one endpoint of \(z\) is sent through \(z_p\) as part of \(l([v], z)\), which is a variant value carrying a tuple. The other endpoint of \(z\) is kept, and used to rebind \([c]\) (using let) with a “new” record, consisting in all the entries of the “original” \([c]\), except \(z_p\) (which has been used for output). More in detail, the “new” \([c]\) has an entry for \(p\) (mapping \(p\) to \(z\)) if \(S'\) still involves \(p\) (otherwise, if \(p \not\in S'\), then \(z\) is discarded, since it has type \([S']_p = \text{end} = \bullet\). After let, the encoding continues as \([P]\).

Symmetrically, branching on \(c[p]\) is also encoded using information from the typing context, i.e., that \(c\) has type \(S = p \&_{\forall i : I} \uplus (U_i).S'_i\) — and therefore, \([S]\) is a record type with one entry \(q : z_q\) for each \(q \in S\). As above, the encoded process deconstructs \([c]\) (using with), and then uses the (linear) name in its \(p\)-entry to perform an input \(z_p(y); y\) is assumed to be a variant, and is pattern matched to determine the continuation. If \(y\) matches \(l_i\) (for some \(i \in I\)), and it carries a tuple \(z_i = (x_i, z)\) (where \(z\) is a continuation name), then \([c]\) is rebound (using let) and the process continues as \([P_i]\). The rebinding of \([c]\) depends on \(l_i\) and the continuation type \(S'_i\): the “new” \([c]\) is a record with all the linear names of the “original” \([c]\), except \(z_p\) (which has been used for input); as above, an entry for \(p\) will exist (and map \(p\) to \(z\)) if \(S'_i\) still involves \(p\) (otherwise, if \(p \not\in S'_i\), then \(z\) has type \(\bullet\) and is discarded.

We will explain the encoding of session restriction \((\nu s)P\) later, after Definition 5.7, as it requires some technicalities: namely, the substitution \(\sigma(\Gamma')\). We can, however, have an intuition about the role of \(\sigma(\Gamma')\) by considering an obvious discrepancy. Consider the following session \(\pi\)-calculus process, that reduces by communication (cf. Definition 2.3):

\[
\Gamma, s[p]:S, s[q]:S' \vdash s[p] \& \{l(x).P\} \mid s[q] \oplus (l(v)).Q \rightarrow P\{/z\} \mid Q
\]  

(1)

We would like its encoding to reduce and communicate, too — but it is not the case:

\[
\text{with } [r : z_r]_{r \in S} = [s[p]] \text{ do } \ldots \text{ with } [r : z_r]_{r \in S'} = [s[q]] \text{ do } \ldots /\!
\]

(2)
The encoding of typing judgements is shown in Figure 7. Here, $\llbracket P \rrbracket_{\Theta, \Gamma} = Q$ iff $\Theta, \Gamma \vdash P$ $\Rightarrow$ $\Theta, \Gamma \vdash Q$ (Definition 5.7).

Reification of Multiparty Sessions. By simply translating a channel with role $s[p]$ into a $\pi$-calculus name $z_{s[p]}$, we have not yet captured the insight behind our approach, i.e., the idea that a multiparty session can be decomposed into a labelled tuple of linear channels (i.e., $\pi$-calculus names), naming pairs of roles. We can formalise “connections” as follows.

Definition 5.5. The connections of $s$ in $\Gamma$ are: $\text{conn}(s, \Gamma) \triangleq \{ (p, q) \mid s[p] ; S_p \in \Gamma \land q \in S_p \}$

Intuitively, two roles $p, q$ are connected by $s$ in $\Gamma$ if $p$ occurs in the type $\Gamma(s[q])$ (but $q$ might not occur in $\Gamma(s[p])$); note, however, that $q$ will always occur if $\Gamma$ is consistent.

Now, as anticipated above, we want to substitute each $s[p]$ with a suitably defined record, containing $\pi$-calculus names; moreover, such names must be typed in the typing context. But what are exactly such names, and their types? This is answered by Definition 5.6.

Definition 5.6 (Reification and decomposition of MPST contexts). The reification of a session typing context $\Gamma_S$ is the substitution:

$$\sigma(\Gamma_S) = \{ [z : (s \cdot p \cdot a)]_{s \cdot S_p} / [s[p]]_{s[p] ; S_p \in \Gamma_S} \}$$

The linear decomposition of $\Gamma_S$ is the $\pi$-calculus typing context $\delta(\Gamma_S)$, defined as:

$$\delta(\Gamma_S) = \bigoplus_{s[p] ; S_p \in \Gamma_S} \{ z_{(s \cdot p \cdot q)} : [\text{unf}(S_p \cdot q)] \}$$
The simplest part of Definition 5.6 is \( \pi \)-calculus reification typing rule is (note that \( \Gamma_5, \Gamma'_5 \) are MPST typing contexts):

\[
\frac{\Theta, \Gamma_5 \vdash P}{\Theta, \Gamma'_5 \vdash \text{T}_\pi \text{-REIFY} \Downarrow \Theta, \Gamma'_5} \]

The definition of \( \sigma(\Gamma_5) \) was the last ingredient needed to formalise our encoding, presented in Definition 5.7 below. The rest of Definition 5.6 will be used later on, to prove its correctness (Theorem 6.2): hence, we postpone its explanation to page 22.

**Definition 5.7 (Encoding).** The encoding of session typing judgements is given in Figure 7. We define \( [P]_{\Theta, \Gamma} = Q \) iff \( \Theta \cdot \Gamma \vdash P = \Theta \cdot \Gamma \vdash Q \). Sometimes, we write \( [P] \) for \( [P]_{\Theta, \Gamma} \) when \( \Theta, \Gamma \) are empty, or clear from the context.

We conclude by explaining the last case in Figure 7, which was not addressed on p.19. The process \( (\mu s, s_s, s_c)Q \mid P_s \mid P_c \mid P_d \) where

\[
Q = (\mu s)\left( s_u(s|p) \oplus (\text{play}(s|a)) \mid s_u(s|q) \oplus (\text{play}(s|b)) \mid s_v(s|p) \oplus (\text{play}(s|c)) \right)
\]

This encoding is the following process, with \( s \) decomposed into 3 linear channels (see Figure 8):

\[
(\mu s)\left( s_u(s|p) \oplus (\text{play}(s|a)) \mid s_u(s|q) \oplus (\text{play}(s|b)) \mid s_v(s|p) \oplus (\text{play}(s|c)) \right)
\]

**Example 5.8.** Consider (1). If we delimit \( s \) and encode the resulting process, we obtain a \( \pi \)-calculus process based on (2), enclosed by the delimitations yielded by \( (\mu s) \), and the substitution \( \sigma(s|p):S, s|q:S', ... \). Since the latter replaces \( \mu_s(p), \mu_s(q) \) with records whose entries reflect roles(\( S \)) and roles(\( S' \)), the encoding can now reduce, firing the two withs.

**Example 5.9.** Consider the main server/clients parallel composition in Example 2.2:

\[
(\mu s, s_s, s_c)(Q \mid P_s \mid P_c \mid P_d)
\]

where

\[
Q = (\mu s)\left( s_u(s|p) \oplus (\text{play}(s|a)) \mid s_u(s|q) \oplus (\text{play}(s|b)) \mid s_v(s|p) \oplus (\text{play}(s|c)) \right)
\]

This encoding is the following process, with \( s \) decomposed into 3 linear channels (see Figure 8):

\[
(\mu s)\left( s_u(s|p) \oplus (\text{play}(s|a)) \mid s_u(s|q) \oplus (\text{play}(s|b)) \mid s_v(s|p) \oplus (\text{play}(s|c)) \right)
\]

# Properties of the Encoding

In this section we present some crucial properties ensuring the correctness of our encoding.
Encoding of Types. Theorem 6.1 below says that our encoding comutes the duality between partial session types (Definition 2.8) and $\pi$-types (Definition 4.1), and also preserves subtyping.

- **Theorem 6.1 (Duality/subtyping preservation).** \([H] = [H'];\) if \(U \leq_{\Sigma} U';\) then \([U] \leq_{\pi} [U']\).

Encoding of Typing Judgements. Theorem 6.2 shows that the encoding of session typing judgements into \(\pi\)-calculus typing judgements is valid. As a consequence, a well-typed MPST process also enjoys the type safety guarantees that can be expressed in standard \(\pi\)-calculus.

- **Theorem 6.2 (Correctness of encoding).** \(\Gamma \vdash v : U\) implies \([\Gamma] \vdash [v][U]\), \(\Theta \vdash X : \hat{U}\) implies \([\Theta] \vdash [X][U]\), and \(\Theta \vdash P \implies [\Theta] \vdash P\).

The proof is by induction on the MPST typing derivation, and yields a corresponding \(\pi\)-calculus typing derivation. One simple case is the following, that relates subtyping:

\[
(\text{T-Sub}) \quad \frac{\Theta \vdash \Gamma, c : U \vdash P \quad U' \subseteq_{\Upsilon} U}{\Theta \vdash \Gamma, c : U' \vdash P}
\]

and holds by the induction hypothesis and Theorem 6.1. The most delicate case is the encoding of session restriction \(\Theta \vdash \Gamma \vdash (\nu s: \Gamma') P\) (Figure 7): its encoding turns \((\nu s)\) into a set of delimited names, used in the substitution \(\sigma(\Gamma')\) applied to \([P]\)\(\Theta, \Gamma, \Gamma\). Hence, to prove Theorem 6.2 in this case, we need to type such names, i.e., produce a context that types \([P]\)\(\Theta, \Gamma, \Gamma\)\(\sigma(\Gamma')\). This is where \(\sigma(\Gamma')\) and \((\text{T-\text{Narrow}})\) (Definition 5.6) come into play, as we now explain.

More on reification and decomposition. By Definition 5.6, the typing context \(\delta(\Gamma'_5)\), when defined, \(\delta(\Gamma'_5)\) has an entry for each role of each channel in \(\Gamma_5\); more precisely, an entry \(z_{[s,p,q]}\) for each \(s[p] : S_p \in \Gamma_5\) and \(q \in S_p\). Such entries are used to type the records yielded by \(\sigma(\Gamma'_5)\). The type of \(z_{[s,p,q]}\) is based on the encoding of the unfolded partial projection \(S_p \upharpoonright q\) that can be either \(*\), or \(\text{Li}(T)/\text{Lo}(T)\) (for some \(T\)). Note that if there is also some \(s[q] : S_q \in \Gamma_5\) with \(p \in S_q\), the type of \(z_{[s,p,q]} = z_{[s,p,q]}\) (when defined) is \([\text{unf}(S_p \upharpoonright q)] \cap [\text{unf}(S_q \upharpoonright p)]\).

This creates a deep correspondence between the consistency of \(\Gamma_5\) and the existence of \(\delta(\Gamma'_5)\), shown in Theorem 6.3: the precondition for MPST type safety (i.e., consistency of \(\Gamma_5\)) is precisely characterised in \(\pi\)-calculus by the linear decomposition at the roots of our encoding.

- **Theorem 6.3 (Precise decomposition).** \(\Gamma_5\) is consistent if and only if \(\delta(\Gamma'_5)\) is defined.

The final part of Definition 5.6 is the \(\pi\)-calculus typing rule \((\text{T-\text{Reify}})\), that uses \(\delta(\Gamma'_5)\) to type a process on which \(\sigma(\Gamma'_5)\) has been applied. Intuitively, \(\delta(\Gamma'_5)\) provides a typing context that types each record yielded by \(\sigma(\Gamma'_5)\). We now explain how the rule works and why it is sound (with a slight simplification). Let \(\Gamma'_5 = \{s[p] : S_p\}_{p \in I}\) for some \(I\). Then, by Definition 5.6:

\[
\delta(\Gamma'_5) = \bigcap_{p \in I} \{ z_{[s,p,q]} : [\text{unf}(S_p \upharpoonright q)] \}_{[s,p,q] \in \text{conn}(s,\Gamma_5)}
\]

\[
\sigma(\Gamma'_5) = \{ [q : z_{[s,p,q] }]_{[s,p,q] \in S_p} \}_{p \in I}
\]

(Note: \(\delta(\Gamma'_5)\) is defined iff \(\Gamma'_5\) is consistent, by Theorem 6.3). Take the I/O types yielded by \(\delta(\Gamma'_5)\), i.e., \(\{ T_{[s,p,q]} : [s,p,q] \}_{[s,p,q] \in \text{conn}(s,\Gamma_5)}\) such that \(\delta(\Gamma'_5) = \bigcup_{p \in I} \{ z_{[s,p,q]} : T_{[s,p,q]} \}_{[s,p,q] \in \text{conn}(s,\Gamma_5)}\) (note \(T_{[s,p,q]} : T_{[s,p,q]}\) are distinct). If we assume \(\Theta \vdash \Gamma, \Gamma_5; P\), this derivation holds:

\[
\begin{align*}
\forall q \in S_p & \quad T_{[s,p,q]} \leq \{ z_{[s,p,q]} \} [s,p,q] \\
& \quad \text{(T-Sub)}
\end{align*}
\]

\[
\begin{align*}
\forall q \in S_p & \quad [z_{[s,p,q]}] \in S_p \quad [q : z_{[s,p,q]}] \in S_q \\
& \quad \text{(T-Rex)}
\end{align*}
\]

\[
\begin{align*}
\forall q \in S_p & \quad [s,p,q] \in S_p \quad [q : z_{[s,p,q]}] \in S_q \\
& \quad \text{(T-MS consc - Figure 6)}
\end{align*}
\]

\[
\begin{align*}
\Theta \vdash \Gamma, \Gamma_5; P & \quad \text{P} \delta(\Gamma'_5) = \Theta \vdash \Gamma, \Gamma_5; P \sigma(\Gamma'_5)
\end{align*}
\]
In particular, the assumptions $T_{(s,p,q)} \leq \pi [S_p | q]$ hold by Lemma 4.7, since each $T_{(s,p,q)}$ is obtained by splitting $\delta(\Gamma'_2)$ (that combines types with $\cap$) using $\psi$. The equivalence in the conclusion holds since $\text{dom}(\Gamma_2) \cap \text{dom}(\delta(\Gamma'_2)) = \emptyset$. Hence: if the $(T\pi\text{-Reify})$ premise $[[\Theta \cdot \Gamma_2], [\Gamma'_2] \vdash P]$ holds, the above derivation holds, proving the conclusion of $(T\pi\text{-Reify})$.

Now, we can finish the proof of Theorem 6.2 for the case $\Theta \cdot \Gamma \vdash (\nu s : \Gamma') P$. Assuming that the judgement holds, we also have $\Theta \cdot \Gamma, \Gamma' \vdash P$ and $\Gamma'$ complete (by the premise of $T\text{-Res}_1$, Figure 4): hence, $\Gamma'$ is consistent, and $\delta(\Gamma')$ is defined (by Theorem 6.3). Assuming that $[[\Theta \cdot \Gamma, \Gamma' \vdash P]$ holds (by the induction hypothesis), we obtain:

$$[[\Theta \cdot \Gamma], [\Gamma'] \vdash [P]_{\Theta \cdot \Gamma, \Gamma'} \text{ (T\pi\text{-Reify})}$$

where $\delta(\Gamma')$ types all the names $z_{(s,p,q)}$ in $\sigma(\Gamma')$, that are also delimited by $[[\nu s]]$. We can conclude by applying $(T\pi\text{-Res}_1)$ to type such delimitations (cf. Figure 5 — this is allowed by the completeness of $\Gamma'$): we get $[[\Theta \cdot \Gamma] \vdash [\nu s][P]_{\Theta \cdot \Gamma, \Gamma'} \sigma(\Gamma')$. i.e., we match Figure 7.

Finally, notice (from Figure 7) that our encoding of processes uses some typing information. In principle, a process could be typed by applying the rules in multiple ways (especially $(T\text{-Sub})$ in Figure 4), and one might wonder whether an MPST process could have multiple encodings. Proposition 6.4 says that this is not the case: the reason is that the only typing information being used is the set of roles in each session type, which does not depend on the typing rule — and is constant w.r.t. subtyping (i.e., $S \leq S'$ implies roles($S$)=roles($S'$)).

**Proposition 6.4 (Uniqueness).** If $\Theta \cdot \Gamma \vdash P$ and $\Theta' \cdot \Gamma' \vdash P$, then $[P]_{\Theta \cdot \Gamma} = [P]_{\Theta' \cdot \Gamma'}$.

**Encoding and Reduction.** One usual way to assess that an encoding is “behaviourally correct” (i.e., a process and its encoding behave “in the same way”) consists in proving operational correspondence. Roughly, it says that the encoding is:

1. **complete**, i.e., any reduction of the original process is simulated by its encoding; and
2. **sound**, i.e., any reduction of the encoded process matches some reduction of the original process.

This is formalised in Theorem 6.5, where $\xrightarrow{\text{with}}$ denotes a reduction induced by $(R\pi\text{-With})$ (Definition 3.2).

**Theorem 6.5 (Operational correspondence).** If $\emptyset \cdot \emptyset \vdash P$, then:

1. (Completeness) $P \rightarrow P''$ implies $\exists \bar{x}, P''$, such that $[P] \rightarrow^{\pi} (\nu \bar{x}) P''$ and $P'' = [P]$;
2. (Soundness) $[P] \rightarrow^{\pi} P'$ implies $\exists \bar{x}, P''', P''$, s.t. $P \rightarrow^{\pi} (\nu \bar{x}) P'''$, $P \rightarrow^{\pi} P''$ and $[P] \xrightarrow{\text{with}} P''$.

The statement of Theorem 6.5 is standard [23, §5.1.3]. Item 1 says that if $P$ reduces to $P'$, then the encoding of the former can reduce to the encoding of the latter. Item 2 says (roughly) that no matter how the encoding of $P$ reduces, it can always further reduce to the encoding of some $P''$, such that $P$ reduces to $P'$. Note that when we write $[P]'$, we mean $[P]'_{S \cdot \emptyset}$, which implies $\emptyset \cdot \emptyset \vdash P'$ (cf. Definition 5.7). The restricted variables $\bar{x}$ in items 1-2 are generated by the encoding of selection (Figure 7): it creates a (delimited) linear name to continue the session. To see why item 2 uses $\xrightarrow{\text{with}}^*$, consider the following MPST process:

$$\emptyset \cdot \Gamma, s[p] : S \vdash s[p][q] & \{(\pi x). P\} \not\rightarrow$$ (the process is stuck)

If we encode it (and apply $\sigma(\Gamma,s[p] : S)$ as per Example 5.8), we get a $\pi$-calculus process that gets stuck, too — but only after firing one internal $\text{with}$-reduction:

$$\text{with} [x : z]_{x \in S} = [x : z_{(s,p,q)}]_{x \in S} \text{ do } z_{(s,p,q)}(y) \ldots \xrightarrow{\text{with}} z_{(s,p,q)}(y) \ldots \not\rightarrow$$

This happens whenever a process is deadlocked, because in Figure 7, the “atomic” MPST branch/select actions are encoded with multiple $\pi$-calculus steps: first $\text{with}$ to deconstruct.
the channels tuple, and then input/output. In general, if an MPST process is stuck, its encoding fires one with for each branch/select, then blocks on an input/output.

Theorem 6.5 yields a corollary on deadlock freedom (Corollary 6.6), that in turn allows to transfer deadlock freedom (Theorem 2.19) from MPST to $\pi$-calculus processes (Corollary 6.7 below).

- **Corollary 6.6.** $P$ is deadlock-free if and only if $\llbracket P \rrbracket$ is deadlock-free, i.e.: $[P] \rightarrow^* P' \not\rightarrow$ implies $\exists Q \equiv 0$ such that $P' = \llbracket Q \rrbracket$.

- **Corollary 6.7.** Let $\llbracket \emptyset \cdot \emptyset \vdash P \rrbracket$, where $P \equiv (\nu s \cdot G)\mid_{i \in I} P_i$ and each $P_i$ only interacts on $s[p_i]$. Then, $\llbracket P \rrbracket$ is deadlock-free.

### 7 From Theory to Implementation

We can now show how our encoding directly guides the implementation of a toolchain for generating safe multiparty session APIs in Scala, supporting distributed delegation. We continue our Game example from Section 1, focusing on player $b$: we sketch the API generation and an implementation of a client, following the results in Section 6. Our approach is to:

1. exploit type safety and distribution provided by an existing library for binary session channels, and then
2. treat the ordering of communications across separate channels in the API generation.

**Scala and lchannels.** Our Scala toolchain is built upon the lchannels library [61, 62]. lchannels provides two key classes, $\text{Out}[T]$ and $\text{In}[T]$, whose instances must be used linearly (i.e., once) to send/receive (by method calls) a $\tau$-typed message: i.e., they represent channel endpoints with $\pi$-calculus types $\text{Lo}(T)$ and $\text{Li}(T)$ (Definition 3.3). This approach enforces the typing of I/O actions via static Scala typing; the linear usage of channels, instead, goes beyond the capabilities of the Scala typing system, and is therefore enforced with run-time checks.

lchannels delivers messages by abstracting over various transports: local memory, TCP sockets, Akka actors [41]. Notably, lchannels promotes session type-safety through a continuation-passing-style encoding of binary session types [61] that is close to our encoding of partial projections (formalised in Definition 5.1). Further, lchannels allows to send/receive $\text{In}[T]/\text{Out}[T]$ instances for binary session delegation [61, Example 4.3]; on distributed message transports, instances of $\text{In}[T]/\text{Out}[T]$ can be sent remotely (e.g., via the Akka-based transport).

**Type-safe, distributed multiparty delegation.** By Theorem 6.2, Definition 5.1 and Theorem 6.3, we know that the game player session type $S_b$ in our example (see Section 1, page 3) provides the type safety guarantees of a tuple of (linear) channels, whose types are given by the encoded partial projections of $S_b$ onto $a$ and $c$ (Definition 2.9). This suggests that, using lchannels, the delegation of an $S_b$-typed channel (as seen in Section 1) could be rendered in Scala as:

```
In[PlayB] with definitions:
  case class PlayB(payload: S_b)
  case class S_b(a: In[InfoAB], c: Out[InfoBC])
```
i.e., as a linear input channel carrying a message of type PlayB, whose payload has type $S_b$; $S_b$, in turn, is a Scala case class, which can be seen as a labelled tuple, that maps $a,c$ to I/O channels — whose types derive from $[S_b][a]$ and $[S_b][c]$ (in fact, they carry messages of type $\text{InfoAB}/\text{InfoBC}$). In this view, $S_b$ is our Scala rendering of the encoded session type $\llbracket S_b \rrbracket$. As said above, lchannels allows to send channels remotely — hence, also allows to remotely
send tuples of channels (e.g., instances of $S_b$); thus, with this simple approach, we obtain type-safe distributed multiparty delegation of an $[S_b]$-typed channel tuple “for free”.

**Multiparty API generation.** Corresponding to the $\pi$-calculus labelled tuple type yielded by the type encoding $[S_b]$, the $S_b$ class outlined above can ensure communication safety, i.e., no unexpected message will be sent or received on any of its binary channels. Like $[S_b]$, however, $S_b$ does not convey any ordering to communications across channels: i.e., $S_b$ does not suggest the order in which its fields $a,c$ should be used. (Indeed, $[S_b]$ may type $\pi$-processes that use its separate channels in any order, while preserving type safety.) To recover the “desired” ordering of communications, and implement it correctly, we can refine our classes so that:

1. each multiparty channel class (e.g., $S_b$) exposes a send() or receive() method, according to the I/O action expected by the multiparty session type (e.g., $S_a$);
2. the implementation of such method uses the binary channels as per our process encoding. E.g., consider again $S_b$ and $S_a$. $S_a$ requires to send towards $c$, so $S_a$ could provide the API:

```scala
case class $S_a$(a: In[InfoAB], c: Out[InfoBC]) {
  def send(v: String) = { // $v$ is the payload of InfoBC message
    val c' = c !! InfoBC(v) // lchannels method: send $v$, and return continuation
    $S'_a$(a, c')
  }
}
```

Now, $S_a$.send() behaves exactly as our process encoding in Figure 7 (case for selection $\oplus$): it picks the correct channel from the tuple (in this case, $c$), creates a new tuple $S'_a$ where $c$ maps to a continuation channel, and returns it — so that the caller can use it to continue the multiparty session interaction. The class $S'_a$ should be similar, with a receive() method that uses a for input (by following the encoding of $\&$). This way, a programmer is correctly led to write, e.g., val $x = s$.send(...).receive() (using method call chaining) — whereas attempting, e.g., s.receive() is rejected by the Scala compiler (method undefined). These send()/receive() APIs are mechanical, and can be automatically generated: we did it by extending Scribble.

**Scribble-Scala Toolchain.** Scribble is a practical MPST-based language and tool for describing global protocols [63, 68]. To implement our results, we have extended Scribble (both the language and the tool) to support the full MPST theory in Section 2, including, e.g., projection, type merging and delegation (not previously supported). Our extension supports protocols with the syntax in Figure 9 (left), by augmenting Scribble with a projection operator $\Phi$; then, it computes the projections/encodings explained in Section 5, and automates the Scala API generation as outlined above (producing, e.g., the $S_a$, $S'_a$,... classes and their send/receive methods). This approach reminds the Java API generation in [29] — but we follow a formal foundation and target the type-safe binary channels provided by lchannels (that, as shown above, takes care of most irksome aspects — e.g., delegation). As a result, the $P_c$ client in Figure 1 can be written as in Figure 9 (right); and although conceptually programmed as Figure 2, the networking mechanisms of the game will concretely follow Figure 8.

**8 Conclusion and Related Works**

We presented the first encoding of a full-fledged multiparty session $\pi$-calculus into standard $\pi$-calculus (Section 5), and used it as the foundation of the first implementation of multiparty sessions (based on Scala API generation) supporting distributed multiparty delegation, on top of existing libraries (Section 7). We proved that the type safety property of MPST is precisely characterised by our decomposition into linear $\pi$-calculus (Theorem 6.3). We encode types by preserving duality and subtyping (Theorem 6.1); our encoding of processes is type-preserving,
A Linear Decomposition of Multiparty Sessions

```scala
global protocol ClientA(role p, role q) {
  PlayA(Game@a) from q to p; // Delegation payload
}
global protocol ClientB(role p, role q) {
  PlayB(Game@b) from q to p; // Delegation payload
}
global protocol ClientC(role p, role q) {
  PlayC(Game@c) from q to p; // Delegation payload
}
global protocol Game(role a, role b, role c) {
  InfoBC(String) from b to c;
  InfoCA(String) from c to a;
  InfoAB(String) from a to b;
  rec t {
    choice at a {
      Mov1AB(Int) from a to b;
      Mov1BC(Int) from b to c;
      choice at c {
        Mov1CA(Int) from c to a; continue t;
      } or {
        Mov2CA(Bool) from c to a; continue t;
      }
    } or {
      Mov2AB(Bool) from a to b;
      Mov2BC(Bool) from b to c;
      choice at c {
        Mov1CA(Int) from c to a; continue t;
      } or {
        Mov2CA(Bool) from c to a; continue t;
      }
    }
  }
}
def P_b(c_bin: In[binary.PlayB]) = { // Cf. Ex.2.2
  // Wrap binary chan in generated multiparty API
  Client_b(MPPlayB(c_bin))
}
def Client_b(y: MPPlayB): Unit = {
  val z = y.receive().p // p is the payload field
  // Send info to c; wait info from a; enter loop
  Loop_b(z.send(InfoBC("...")).receive())
}
def Loop_b(x: MPMov1ABOrMov2AB): Unit = {
  x.receive() match {
    case Mov1AB(p, cont) => {
      // cont only allows to send Mov1BC
      Loop_b(cont.send(Mov1BC(p)));
    }
    case Mov2AB(p, cont) => {
      // cont only allows to send Mov2BC
      Loop_b(cont.send(Mov2BC(p)));
    }
  }}
}
```

Figure 9 Game example (Section 1). Left: Scribble protocols for client/server setup sessions, and main Game (Example 2.18). Right: Scala code for player b, using Scribble-generated APIs to mimick Example 2.2.

and operationally sound and complete (Theorem 6.2 and Theorem 6.5); hence, our encoding preserves the type-safety and deadlock-freedom properties of MPST (Corollary 6.7). These results ensure the correctness of our (encoding-based) Scala implementation. Moreover, our encoding preserves process distribution (i.e., is homomorphic w.r.t. parallel composition); correspondingly, our implementation of multiparty sessions is decentralised and choreographic.

Session Types for “Mainstream” Languages. We mentioned binary session implementations for various languages in Section 1. Notably, [57, 32, 33, 42, 52, 61, 55] seek to integrate session types in the native host language, without language extensions, to avoid hindering their use in practice. To do so, one approach (e.g. in [61, 55]) is combining static typing of I/O actions with run-time checking of linear channel usage. Our implementation adopts this idea (Section 7). Haskell-based works exploit its richer typing system to statically enforce linearity — with various expressiveness/usability trade-offs based on their session types embedding strategy.

Implementations of multiparty sessions are few and limited, due to the intricacies of the theory (e.g., the interplay between projections, mergability and consistency), and practical issues (e.g., realising multiparty abstractions over binary transports, including distributed delegation), as discussed in Section 1. [64] was the first implementation of MPST, based on extending Java with session primitives. [29] proposes MPST-based API generation for Java, based on CFMSs [7], but has no formalisation — unlike our implementation, that follows our encoding. [17, 20] develop MPST-influenced networking APIs in Python and Erlang; [50] implements recovery strategies in Erlang. [17, 20, 50] focus on purely dynamic MPST verification via run-time monitoring. [51, 48] extends [17] with actors and timed specifications. [46] uses a dependent MPST theory to verify MPI programs. Crucially, none of these implementations supports delegation (nor projection merging, needed by our Game example, cf. Example 2.14).

Encodings of Session Types and Processes. [16] encodes binary session $\pi$-calculus into an augmented $\pi$-calculus with branch/select constructs. [15], following [36], and [21] encode
non-recursive, binary session π-calculus, respectively into linear π-calculus and the Generic Type System for π-calculus [31], proving correctness w.r.t. typing and reduction. All the above works investigate binary and (except [16]) non-recursive session types, while in this paper we study the encoding of multiparty session types, subsuming binary ones; and unlike [16], we target standard π-calculus. We encode branching/selection using variants as in [15, 13], but our treatment of recursion, and the rest of the MPST theory, is novel.

Encodings of multiparty into binary sessions are studied in [9, 8]. Both use orchestration approaches that add centralised medium/arbiter processes, and target session calculi (not π-calculus). [53] uses a limited class of global types to extract “characteristic” deadlock-free π-calculus processes — without addressing session calculi, nor proving operational properties.

Recursion and Duality. The interplay between recursion and duality has been a thorny issue in session types literature, requiring our careful treatment in Section 4. [6, 1] noticed that the standard duality in [26] does not commute with the unfolding of recursion when type variables occur as payload, e.g., µt.τ.end. To solve this issue, [6, 1] define a new notion of duality, called complement [1], then used in [13] to encode recursive binary session types into linear π-types. Unfortunately, [2] later noticed that even complement does not commute, e.g., when unfolding µt.µt′.τ.t′. As observed in Section 4, to encode recursive session types we encounter similar issues in π-types. The reason seems natural: τ-types do not distinguish “payloads” and “continuations”, and in recursive linear inputs/outputs, type variables always occur as “payload”, e.g., µt.Lo(t). Since, in the light of [2], we could not adopt the approach of [13], we propose a solution similar to [43]: introduce dualised type variables t. [43] also sketches a property similar to our Lemma 4.4. The main difference is that we add dualised variables to π-types (while [43] adds t to session types). An alternative idea is given in [61]: encoding recursive session types as non-recursive linear I/O types with recursive payloads. This avoids dualised variables (e.g., Lo/µt.Li(t)) instead of µt.Lo(t), but if adopted, would complicate Definition 5.1. Moreover, [61] studies the encoding of recursive types, but not processes.

Future work. On the practical side, we plan to study whether Scala language extensions could provide stronger static channel usage checks. E.g., [25, 24] (capabilities) could allow to ensure that a channel endpoint is not used after being sent; [58, 65] (effects) could allow to ensure that a channel endpoint is actually used in a given method. We also plan to extend our multiparty API generation approach beyond Scala and channels, targeting other languages and implementations of binary sessions/channels [57, 32, 33, 42, 52, 55].

On the theoretical side, our encoding provides a basis for reusing theoretical results and tools between MPST π-calculus and standard π-calculus. E.g., we could now exploit Corollary 6.6, to verify deadlock-freedom of processes with interleaved multiparty sessions (studied in [3, 10, 12]) by applying π-calculus deadlock detection methods to their encodings [38, 35, 66]. Moreover, we can prove that our encoding is barb-preserving: hence, we plan to study its full abstraction properties w.r.t. barbed congruence in session π-calculus [40, 39] and π-calculus.

Thanks to the reviewers for their remarks, and to B. Toninho for fruitful discussions. Thanks to S.S. Jongmans, R. Neykova, N. Ng, B. Toninho for testing the companion artifact.
References


Simon Fowler. An Erlang implementation of multiparty session actors. In ICE, 2016. doi:10.4204/EPTCS.223.3.


A Linear Decomposition of Multiparty Sessions


53 Luca Padovani. Deadlock and lock freedom in the linear π-calculus. Online version of [54], January 2014. URL: https://hal.inria.fr/hal-00932356.


TYPICAL. Type-based static analyzer for the pi-calculus. http://www-kb.is.s.u-tokyo.ac.jp/~koba/typical/.
