Exceptional Asynchronous Session Types

Session Types without Tiers

SIMON FOWLER, The University of Edinburgh
SAM LINDLEY, The University of Edinburgh
J. GARRETT MORRIS, University of Kansas
SÁRA DECOVA, The University of Edinburgh

Session types statically guarantee that communication complies with a protocol. However, most accounts of session typing do not account for failure, which means they are of limited use in real applications—especially distributed applications—where failure is pervasive.

We present the first formal integration of asynchronous session types with exception handling in a functional programming language. We define a core calculus which satisfies preservation and progress properties, is deadlock free, confluent, and terminating.

We provide the first implementation of session types with exception handling for a fully-fledged functional programming language, by extending the Links web programming language; our implementation draws on existing work on effect handlers. We illustrate our approach through a running example of two-factor authentication, and a larger example of a session-based chat application where communication occurs over session-typed channels and disconnections are handled gracefully.

ACM Reference Format:

1 INTRODUCTION

With the growth of the internet and mobile devices, as well as the failure of Moore’s law, concurrency and distribution have become central to many applications. Writing correct concurrent and distributed code requires effective tools for reasoning about communication protocols. While data types provide an effective tool for reasoning about the shape of data communicated, protocols also require us to reason about the order in which messages are transmitted.

Session types [Honda 1993; Honda et al. 1998] are types for protocols. They describe both the shape and order of messages. If a program type-checks according to its session type, then it is statically guaranteed to comply with the corresponding protocol. Alas, most accounts of session types do not handle failure, which means they are of limited use in distributed settings where failure is pervasive. Inspired by work of Mostrous and Vasconcelos [2014], we present the first account of asynchronous session types in a functional programming language, which smoothly handles both distribution and failure. We present both a core calculus enjoying strong...
TwoFactorServer ≜
(?(Username, Password) ⊕
Authenticated : ServerBody,
Challenge : !ChallengeKey.?Response.
⊕{Authenticated : ServerBody,
AccessDenied : End},
AccessDenied : End)

(a) Server Session Type

TwoFactorClient ≜
!(Username, Password) &
{Authenticated : ClientBody,
&{Authenticated : ClientBody,
AccessDenied : End},
AccessDenied : End}

(b) Client Session Type

Fig. 1. Two-factor Authentication Session Types

metatheoretical correctness properties and a practical implementation as an extension of the Links web programming language [Cooper et al. 2007].

1.1 Session Types

We illustrate session types with a basic example of two-factor authentication. A user inputs their credentials. If the login attempt is from a known device, then they are authenticated and may proceed to perform privileged actions. If the login attempt is from an unrecognised device, then the user is sent a challenge code. They enter the challenge code into a hardware key which yields a response code. If the user responds with the correct response code, then they are authenticated.

A session type specifies the communication behaviour of one endpoint of a communication channel participating in a dialogue (or session) with the other endpoint of the channel. Fig. 1 shows the session types of two channel endpoints connecting a client and a server. Fig. 1a shows the session type for the server which first receives (?) a pair of a username and password from the client. Next, the server selects (⊕) whether to authenticate the client, issue a challenge, or reject the credentials. If the server decides to issue a challenge, then it sends (!) the challenge string, awaits the response, and either authenticates or rejects the client. The ServerBody type abstracts over the remainder of the interactions, for example making a deposit or withdrawal.

Duality. The client implements the dual session type, shown in Fig. 1b. Whenever the server receives a value, the client sends a value, and vice versa. Whenever the server makes a selection, the client offers a choice (&), and vice versa. This duality between client and server ensures that each communication is matched by the other party. We denote duality with an overbar; thus TwoFactorClient = \overline{TwoFactorServer} and TwoFactorServer = \overline{TwoFactorClient}.

Implementing Two-factor Authentication. Let us suppose we have constructs for sending and receiving along, and for closing, an endpoint:

\text{send} M N : S \quad \text{where } M \text{ has type } A, \text{ and } N \text{ is an endpoint with session type } !A.S

\text{receive} M : (A \times S) \quad \text{where } M \text{ is an endpoint with session type } ?A.S

\text{close} M : 1 \quad \text{where } M \text{ is an endpoint with session type End}

Let us also suppose we have constructs for selecting and offering a choice:

\text{select } \ell_j M : S_j \quad \text{where } M \text{ is an endpoint with session type } \oplus\{\ell_i : S_i\}_{i \in I}, \text{ and } j \in I

\text{offer} M \{\ell_i(x_i) \mapsto N_i\}_{i \in I} : A \quad \text{where } M \text{ is an endpoint with session type } \&\{\ell_i : S_i\}_{i \in I}, \text{ each } x_i \text{ binds an endpoint with session type } S_i, \text{ and each } N_i \text{ has type } A
We can now write a client implementation.

twoFactorClient : (Username × Password × TwoFactorClient) → 1

twoFactorClient(username, password, s) ≜
  let s = send (username, password) s in
  offer s {Authenticated(s) ↦ clientBody(s)}

  Challenge(s) ↦ let (key, s) = receive s in
  let s = send (generateResponse(key)) s in
  offer s {Authenticated(s) ↦ clientBody(s)}

  AccessDenied(s) ↦ close s; loginFailed

The twoFactorClient function takes the credentials and an endpoint of type TwoFactorClient as its arguments. The credentials are sent along the endpoint, then three choices are offered depending on whether the server authenticates the user, sends a two-factor challenge, or rejects the authentication attempt. If the server authenticates the user, then the program progresses to the main application (clientBody(s)). If the server sends a challenge, then the client receives the challenge key, and sends the response, calculated by generateResponse. Two choices are then offered according to whether the challenge response was successful. The rejection of an authentication attempt is part of the protocol and not exceptional behaviour. We can also write a server implementation.

twoFactorServer : TwoFactorServer → 1

twoFactorServer(s) ≜ let ((username, password), s) = receive s in
  if checkDetails(username, password) then
    let s = select Authenticated s in serverBody(s)
  else
    let s = select AccessDenied s in close s

The twoFactorServer function takes an endpoint of type TwoFactorServer along which it receives the credentials, which are checked using checkDetails. If the check passes, then the server proceeds to the application body (serverBody(s)); if not, then the server notifies the client by selecting the AccessDenied branch. This particular server implementation opts to never send a challenge request.

Statically checking session types demands a substructural type system. We discuss three options: linear types, affine types, and linear types with explicit cancellation.

1.2 Linear Types

Simply providing constructs for sending and receiving values, and for selecting and offering choices, is insufficient for safely implementing session types. Consider the following client:

wrongClient : TwoFactorClient → 1

wrongClient(s) ≜ let t = send ("Alice", "hunter2") s in
  let t = send ("Bob", "letmein") s in ...

Reuse of s allows a (username, password) pair to be sent along the same endpoint twice, violating the fundamental property of session fidelity, which states that in a well-typed program, communication over an endpoint matches its session type. To maintain session fidelity and ensure that all communication actions in a session type occur, session type systems typically require that each endpoint is used linearly—exactly once.

Exceptions In practice, linear session types are unrealistic. Thus far, we have assumed checkDetails always succeeds, which may be plausible if checking against an in-memory store, but not if connecting to a remote database. One option would be for checkDetails to return false on
As a first attempt, we might try to write:

```latex
exnServer1 : TwoFactorClient → 1
exnServer1(s) ≝ let ((username, password), s) = receive s in
  try if checkDetails(username, password) then
    let s = select Authenticated s in serverBody(s)
  else
    let s = select AccessDenied s in close s
  catch log("Database Error")
```

However, the above code does not type-check and is unsafe. Linear endpoint s is not used in the `catch` block and yet is still open if an exception is raised by `checkDetails`.

As a second attempt, we may decide to localise exception handling to the call to `checkDetails`. We introduce `checkDetailsOpt`, which returns `Some(result)` if the call is successful and `None` if not.

```latex
checkDetailsOpt : (Username × Password) → Option(Bool)
checkDetailsOpt(username, password) ≝ try Some(checkDetails(username, password))
```

```latex
exnServer2 : TwoFactorServer → 1
exnServer2(s) ≝ let ((username, password), s) = receive s in
  case checkDetailsOpt(username, password) of
    Some(res) → if res then let s = select Authenticated s in serverBody(s)
    else let s = select AccessDenied s in close s
    None → log("Database Error")
```

Still the code is unsafe as it does not use s in the None branch of the case-split. However, we do now have more precise information about the type of s, since it is unused in the `try` block. One solution could be to adapt the protocol by adding an `InternalError` branch:

```latex
TwoFactorServerExn ≝ ?{(Username, Password).@$\{$
  Authenticated : ServerBody,
  AccessDenied : End,
  InternalError : End$}
```

We could use `select InternalError s` in the None branch to yield a type-correct program, but doing so would be unsatisfactory as it clutters the protocol and the implementation with failure points.

Disconnection. The problem of failure is compounded by the possibility of disconnection. On a single machine it may be plausible to assume that communication always succeeds. In a distributed setting this assumption is unrealistic as parties may disconnect without warning. The problem is particularly acute in web applications as a client may close the browser at any point. In order to adequately handle failure we must incorporate some mechanism for detecting disconnection.

### 1.3 Affine Types

We began by assuming linear types—each endpoint must be used exactly once. One might consider relaxing linear types to affine types—each endpoint must be used at most once. Statically checked affine types form the basis of the existing Rust implementation of session types [Jespersen et al. 2015] and dynamically checked affine types form the basis of the OCaml FuSe [Padovani 2017] and Scala lchannels [Scalas and Yoshida 2016] session type libraries. Affine types present two
quandaries arising from endpoints being silently discarded. First, a developer receives no feedback if they accidentally forget to finish a protocol implementation. Second, if an exception is raised in an evaluation context that captures an open endpoint then the peer may be left waiting forever.

1.4 Linear Types with Explicit Cancellation

Mostrous and Vasconcelos [2014] address the difficulties outlined above through an explicit discard (or cancellation) operator. (They characterise their sessions as affine, but it is important not to confuse their system with affine type systems, as in §1.3, which allow variables to be discarded implicitly.) Their approach boils down to three key principles: endpoints can be explicitly discarded; an exception is thrown if a communication cannot succeed because a peer endpoint has been cancelled; and endpoint cancellations are propagated when endpoints become inaccessible due to an exception being thrown. They introduce a process calculus including the term \(a^\perp\) (“cancel \(a\)”), which indicates that endpoint \(a\) may no longer be used to perform communications. They provide an exception handling construct which attempts a communication action, running an exception handler if the action fails, and show that explicit cancellation is well-behaved: their calculus satisfies preservation and global progress (well-typed processes never get stuck), and is confluent.

Explicit cancellation neatly handles failure while ruling out accidentally incomplete implementations and providing a mechanism for notifying peers when an exception is raised. In this paper we take advantage of explicit cancellation to formalise and implement asynchronous session types with failure handling in a distributed functional programming language; this is not merely a routine adaptation of the ideas of Mostrous and Vasconcelos for the following reasons:

- They present a process calculus, but we work in a functional programming language.
- Communication in their system is synchronous, depending on a rendezvous between sender and receiver. We require asynchronous communication, which is more amenable to implementation in a distributed setting.
- Their exception handling construct is over a single communication action and does not allow nested exception handling. This design is difficult to reconcile with a functional language, as it is inherently non-compositional. Our exception handling construct is compositional.

We define a core concurrent \(\lambda\)-calculus, Exceptional GV (EGV), with asynchronous session-typed communication and exception handling. As with the calculus of Mostrous and Vasconcelos, an exception is raised when a communication action fails. But our compositional exception handling construct can be arbitrarily nested, and allows exception handling over multiple communication actions. Using EGV, we may implement the two factor authentication server as follows:

\[
\text{exnServer3} : \text{TwoFactorServer} \to 1
\]

\[
\text{exnServer3}(s) \triangleq \begin{cases} \text{let } ((\text{username, password}), s) = \text{receive } s \text{ in} & \\
\text{try checkDetails(\text{username, password}) as res in} & \\
\text{if res then let } s = \text{select Authenticated } s \text{ in serverBody}(s) & \\
\text{else let } s = \text{select AccessDenied } s \text{ in close } s & \\
\text{otherwise} & \\
\text{cancel } s; \text{ log("Database Error")} & \end{cases}
\]

Following Benton and Kennedy [2001], an exception handler \(\text{try } L \text{ as } x \text{ in } M \text{ otherwise } N\) takes an explicit success continuation \(M\) as well as the usual failure continuation \(N\). If checkDetails fails with an exception, then \(s\) is safely discarded using \(\text{cancel}\), which takes an endpoint and returns the unit value. Disconnection is handled by cancelling all endpoints associated with a client. If a peer tries to read along a cancelled endpoint then an exception is thrown.
```plaintext
try
  let s = fork (λt. cancel t) in
  let (res, s) = receive s in
  close s; res
as res in
  print ("Result: "+ res)
otherwise print "Error!"

let s = fork (λt.
  let (res, t) = receive t in
  close t; res) in
let f = (λx. send x s) in
raise;

let f = (λx. send x s) in
raise;

let u = fork (λv. cancel v) in
let u = send s u in
close u

(a) Cancellation and Exceptions
(b) Delegation
(c) Closures
```

Fig. 2. Failure Examples

We implement the constructs described by EGV as an extension to Links [Cooper et al. 2007], a functional programming language for the web. Our implementation is based on a minimal translation to effect handlers [Plotkin and Pretnar 2013].

1.5 Contributions

This paper makes five main contributions:

1. Exceptional GV (§2), a core linear lambda calculus extended with asynchronous session-typed channels and exception handling. We prove (§3) that the core calculus enjoys preservation, progress, a strong form of confluence called the diamond property, and termination.

2. Extensions to EGV supporting exception payloads, unrestricted types, and access points (§4).

3. The design and implementation of an extension of the Links web programming language to support tierless web applications which can communicate using session-typed channels (§5).

4. Client and server backends for Links implementing session typing with exception handling (§5.4), drawing on connections with effect handlers [Plotkin and Pretnar 2013].

5. Example applications using the infrastructure (§6). In addition to our two-factor authentication workflow we outline the implementation of a chat server.


The rest of the paper is structured as follows: §2 presents Exceptional GV and §3 its metatheory; §4 discusses extensions to Exceptional GV; §5 describes the implementation; §6 presents a chat application written in Links; §7 discusses related work; and §8 concludes.

2 EXCEPTIONAL GV

In this section, we introduce Exceptional GV (henceforth EGV). GV is a core session-typed linear λ-calculus that has a tight correspondence with classical linear logic [Lindley and Morris 2015; Wadler 2014]. EGV is an asynchronous variant of GV with support for failure handling.

Due to GV’s close correspondence with classical linear logic, EGV has a strong metatheory, enjoying preservation, global progress, the diamond property, and termination. Much like the simply-typed λ-calculus, this well-behaved core must be extended to be expressive enough to write larger applications. Nonetheless, the core calculus alone is expressive enough to support our two-factor authentication example, and to support server applications which gracefully handle disconnection. In §3, we show that cancellation is well-behaved, and does not violate any of the core properties of GV. In §4, following Lindley and Morris [2015, 2017], we extend EGV modularly.
with standard features of our implementation, some of which provide weaker guarantees. Channel cancellation and exceptions are orthogonal to these features.

2.1 Integrating Sessions with Exceptions, by Example

Integrating session types with failure handling into a higher-order functional language requires care. Fig. 2 illustrates three important cases: cancellation and exceptions, delegation, and closures. In order to initiate a session, we adopt the fork primitive of Lindley and Morris [2015]. Given a term \( M \) of type \( S \rightarrow 1 \), the term \( \text{fork} M \) of type \( S \) creates a fresh channel with endpoints \( a \) of type \( S \) and \( b \) of type \( \bar{S} \), forks a child thread that executes \( M \ a \), and returns endpoint \( b \).

Cancellation and Exceptions. Fig. 2a forks a thread which immediately cancels its endpoint. The parent attempts to receive, but the message can never arrive so an exception is raised and the otherwise clause is invoked.

Delegation. A central feature of \( \pi \)-calculus is mobility of names. In session calculi sending an endpoint is known as session delegation. The code in Fig. 2b begins by forking a thread and returning endpoint \( s \). The child is passed endpoint \( t \) on which it blocks receiving. Next, the parent forks a second child, yielding endpoint \( u \). The second child is passed endpoint \( v \), which is immediately discarded using cancel. Now the parent thread sends endpoint \( s \) along \( u \). Endpoint \( s \) will never be received as the peer endpoint \( v \) of \( u \) has been cancelled. In turn, this renders \( s \) irretrievable and an exception is thrown in the first child thread, as it can never receive a value.

Closures. It is crucial that cancellation plays nicely with closures. The code in Fig. 2c defines a function \( f \) which sends its argument \( x \) along \( s \). The parent thread then raises an exception. As \( s \) appears in the closure bound to \( f \), which appears in the continuation and is thus discarded, \( s \) must be cancelled.

2.2 Syntax and Typing Rules for Terms

Fig. 3 gives the syntax of EGV. Types include unit (1), linear functions \( (A \rightarrow B) \), linear sums \( (A + B) \), linear tensor products \( (A \times B) \), and session types \( (S) \).

Terms include variables \( (x) \) and the usual introduction and elimination forms for linear functions, unit, products, and sums. We write \( M; N \) as syntactic sugar for \( \text{let} () = M \ in \ N \) and \( \text{let} x = M \ in \ N \) for \( (\lambda x. N) \ M \). The standard session typing primitives [Lindley and Morris 2015] are as follows: fork \( M \) creates a fresh channel with endpoints \( a \) of type \( S \) and \( b \) of type \( \bar{S} \), forks a child thread that executes \( M \ a \), and returns endpoint \( b \); send \( M \ N \) sends \( M \) along endpoint \( N \); receive \( M \) receives along endpoint \( M \); and close \( M \) closes an endpoint when a session is complete.
We introduce three new term constructs to support session typing with failure handling: \texttt{cancel} \( M \) explicitly discards session endpoint \( M \); \texttt{raise} raises an exception; and \texttt{try \( L \) as \( x \) in \( M \) otherwise \( N \)} evaluates \( L \), on success binding the result to \( x \) in \( M \) and on failure evaluating \( N \).

\textit{Explicit success continuations.} Benton and Kennedy [2001] argue that:

From the points of view of programming pragmatics, rewriting and operational semantics, the syntactic construct used for exception handling in ML-like programming languages, and in much theoretical work on exceptions, has subtly undesirable features.

Benton and Kennedy show that explicit success continuations avoid the subtly undesirable features they identify; correspondingly, we adopt their construct. Moreover, explicit success continuations align with the definition of handlers for algebraic effects [Plotkin and Pretnar 2013] that we use in our implementation (§5.4).

\textit{Branching and selection.} Though our implementation supports \texttt{select} and \texttt{offer} directly, and we use them in examples, we omit them from the core calculus (following Lindley and Morris [2015, 2017]) as they can be encoded using sums and delegation [Dardha et al. 2017; Kobayashi 2003].

\textit{Typing.} Fig. 4 gives the typing rules for EGV. As usual, linearity is enforced by splitting environments when typing subterms, ensuring T-Var takes a singleton environment, and leaf rules T-Unit and T-Raise take an empty environment. We write \( \Gamma_1, \Gamma_2 \) to mean the disjoint union of \( \Gamma_1 \) and \( \Gamma_2 \).

The bulk of the rules are standard for a linear \( \lambda \)-calculus. Session types are related by \textit{duality}. The T-Fork rule forks a thread connected by dual endpoints of a channel. The rules T-Send, T-Recev, and T-Close capture session-typed communication.

![Fig. 4. Term Typing and Duality](image-url)

---

### Term Typing

\begin{align*}
\text{T-Var} & : \quad \Gamma, x : A \vdash M : B \\
\text{T-App} & : \quad \Gamma_1 \vdash M : A \to B \quad \Gamma_2 \vdash N : A \\
\text{T-LetUnit} & : \quad \Gamma_1 \vdash M : 1 \\
\text{T-Pair} & : \quad \Gamma_1 \vdash M : A \times B \\
\text{T-Case} & : \quad \Gamma_1 \vdash M : (A \times B) \\
\text{T-Cancel} & : \quad \Gamma \vdash \text{cancel} M : 1 \\
\text{T-Recv} & : \quad \Gamma \vdash \text{recv} M : (A \times B) \\
\text{T-Try} & : \quad \Gamma \vdash \text{try} L \text{ as } x \text{ in } M \text{ otherwise } N : B \\
\text{T-Raise} & : \quad \Gamma \vdash \text{raise} A
\end{align*}

\[ \Gamma \vdash \text{End} \]

---

We introduce three new term constructs to support session typing with failure handling: \texttt{cancel} \( M \) explicitly discards session endpoint \( M \); \texttt{raise} raises an exception; and \texttt{try \( L \) as \( x \) in \( M \) otherwise \( N \)} evaluates \( L \), on success binding the result to \( x \) in \( M \) and on failure evaluating \( N \).

\textit{Explicit success continuations.} Benton and Kennedy [2001] argue that:

From the points of view of programming pragmatics, rewriting and operational semantics, the syntactic construct used for exception handling in ML-like programming languages, and in much theoretical work on exceptions, has subtly undesirable features.

Benton and Kennedy show that explicit success continuations avoid the subtly undesirable features they identify; correspondingly, we adopt their construct. Moreover, explicit success continuations align with the definition of handlers for algebraic effects [Plotkin and Pretnar 2013] that we use in our implementation (§5.4).

\textit{Branching and selection.} Though our implementation supports \texttt{select} and \texttt{offer} directly, and we use them in examples, we omit them from the core calculus (following Lindley and Morris [2015, 2017]) as they can be encoded using sums and delegation [Dardha et al. 2017; Kobayashi 2003].

\textit{Typing.} Fig. 4 gives the typing rules for EGV. As usual, linearity is enforced by splitting environments when typing subterms, ensuring T-Var takes a singleton environment, and leaf rules T-Unit and T-Raise take an empty environment. We write \( \Gamma_1, \Gamma_2 \) to mean the disjoint union of \( \Gamma_1 \) and \( \Gamma_2 \).

The bulk of the rules are standard for a linear \( \lambda \)-calculus. Session types are related by \textit{duality}. The T-Fork rule forks a thread connected by dual endpoints of a channel. The rules T-Send, T-Recev, and T-Close capture session-typed communication.
We now give a small-step operational semantics for EGV.

Runtime Types

\[ R ::= S | S^\# \]

Names

\[ a, b, c \]

Terms

\[ M ::= \cdots | a \]

Values

\[ U, V, W ::= a | \lambda x.M | () | (V, W) | \text{inl} V | \text{inr} V \]

Configurations

\[ C, D, E ::= (va)C | C \parallel D | \phi M | \text{halt} | \xi a | a(V) \rightarrow b(W) \]

Thread Flags

\[ \phi ::= \bullet | \circ \]

Top-level threads

\[ T ::= \bullet M \parallel \text{halt} \]

Auxiliary threads

\[ A ::= o M | \xi a | a(V) \rightarrow b(W) \]

Type Environments

\[ \Gamma ::= \cdots | \Gamma, a : S \]

Runtime Type Environments

\[ \Delta ::= \cdot | \Delta, a : R \]

Evaluation Contexts

\[ E ::= [] | E M | V E \]

\[ \text{let } () = E \text{ in } M | (E, M) | (V, E) | \text{let } (x, y) = E \text{ in } M \]

\[ \text{in } E | \text{inr } E | \text{case } E \text{ of } \{\text{inl } x \mapsto M ; \text{inr } x \mapsto N\} \]

\[ \text{fork } E | \text{send } E M | \text{send } V E | \text{receive } E | \text{close } E \]

\[ \text{cancel } E | \text{try } E \text{ as } x \text{ in } M \text{ otherwise } N \]

Pure Contexts

\[ P ::= [] | P M | V P \]

\[ \text{let } () = P \text{ in } M | \text{let } (x, y) = P \text{ in } M | (P, M) | (V, P) \]

\[ \text{in } P | \text{inr } P | \text{case } P \text{ of } \{\text{inl } x \mapsto M ; \text{inr } x \mapsto N\} \]

\[ \text{fork } P | \text{send } P M | \text{send } V P | \text{receive } P | \text{close } P \]

\[ \text{cancel } P \]

Thread Contexts

\[ \mathcal{T} ::= \phi E \]

Configuration Contexts

\[ \mathcal{G} ::= [] | (va)\mathcal{G} | \mathcal{G} \parallel C \]

Syntactic Sugar

\[ \xi V \triangleq \xi a_1 \parallel \cdots \parallel \xi a_n \text{ where } fn(V) = \{a_i\} \]

\[ \xi P \triangleq \xi a_1 \parallel \cdots \parallel \xi a_n \text{ where } fn(P) = \{a_i\} \]

\[ \xi E \triangleq \xi a_1 \parallel \cdots \parallel \xi a_n \text{ where } fn(E) = \{a_i\} \]

---

As exceptions do not return values, the rule T-RAISE allows an exception to be given any type \(A\). Rule T-TRY embraces explicit success continuations as advocated by Benton and Kennedy [2001], binding a result in \(M\) if \(L\) evaluates successfully. The T-CANCEL rule explicitly discards an endpoint. Naively implemented, cancellation violates progress: a thread could discard an endpoint, leaving a peer waiting forever. We avoid this pitfall by raising an exception when a communication action would wait forever due to cancellation.

2.3 Operational Semantics

We now give a small-step operational semantics for EGV.

**Runtime Syntax.** Fig. 5 shows the runtime syntax of EGV. We write \(S^\#\) for the type of a channel which can be split into two endpoints of types \(S\) and \(\overline{S}\). Runtime types \(R\) are either session types or channel types. We extend the syntax of terms to include names ranged over by \(a, b, c\). Depending on context, a name \(a\) is variously used to identify a channel of type \(S^\#\) and each of its endpoints of type \(S\) and \(\overline{S}\). Values are standard. The semantics makes use of configurations, which are similar to \(\pi\)-calculus processes: \((va)C\) binds name \(a\) in configuration \(C\), and \(C \parallel D\) is the parallel composition of configurations \(C\) and \(D\). Program threads take the form \(\phi M\), where \(\phi\) is a thread flag identifying whether the term is the main thread \((\bullet)\), which returns a top-level result, or a child thread \((\circ)\), which
Term Reduction

\[
E\text{-LAM} \quad (\lambda x.M) V \rightarrow^E M\{V/x\}
\]

\[
E\text{-UNIT} \quad \text{let } (\_ ) = (\_ ) \text{ in } M \rightarrow^E M
\]

\[
E\text{-PAIR} \quad \text{let } (x, y) = (V, W) \text{ in } M \rightarrow^E M\{V/x, W/y\}
\]

\[
E\text{-INL} \quad \text{case inl } V \text{ of } (\text{inl } x \mapsto M; \text{inr } y \mapsto N) \rightarrow^E M\{V/x\}
\]

\[
E\text{-INR} \quad \text{case inr } V \text{ of } (\text{inl } x \mapsto M; \text{inr } y \mapsto N) \rightarrow^E N\{V/y\}
\]

\[
E\text{-VAL} \quad \text{try } V \text{ as } x \text{ in } M \text{ otherwise } N \rightarrow^E M\{V/x\}
\]

\[
E\text{-LIFT} \quad E[M] \rightarrow^E E[M'], \quad \text{if } M \rightarrow^E M'
\]

Configuration Equivalence

\[
C \parallel (D \parallel E) \equiv (C \parallel D) \parallel E \quad \text{if } C \parallel D \equiv D \parallel C
\]

\[
(va)(vb)C \equiv (vb)(va)C
\]

\[
a\Apollo\parallel b\Apollo \equiv b\Apollo \parallel a\Apollo \quad \circ () \equiv C \equiv C
\]

\[
C \parallel (va)D \equiv (va)(C \parallel D), \quad \text{if } a \notin \text{fn}(C)
\]

\[
a\Apollo \parallel b\Apollo \equiv b\Apollo \parallel a\Apollo \quad \circ () \equiv C \equiv C
\]

Configuration Reduction

\[
E\text{-FORK} \quad F[\text{fork } (\lambda x.M)] \rightarrow (va)(vb)(F[a] \parallel \circ M\{b/x\} \parallel (a\Apollo \parallel b\Apollo)) \quad \text{where } a, b \text{ are fresh}
\]

\[
E\text{-SEND} \quad F[\text{send } U] \parallel a\Apollo \parallel b\Apollo \rightarrow F[a] \parallel a\Apollo \parallel b\Apollo \parallel U
\]

\[
E\text{-RECEIVE} \quad F[\text{receive } a] \parallel a\Apollo \parallel b\Apollo \rightarrow F[U, a] \parallel a\Apollo \parallel b\Apollo
\]

\[
E\text{-CLOSE} \quad (va)(vb)(F[\text{close } a] \parallel F'[\text{close } b] \parallel a\Apollo \parallel b\Apollo) \rightarrow F'[U] \parallel F'[0]
\]

\[
E\text{-CANCEL} \quad F[\text{cancel } a] \rightarrow F[U] \parallel \circ a
\]

\[
E\text{-ZAP} \quad \circ a \parallel a\Apollo \parallel b\Apollo \rightarrow \circ a \parallel U \parallel a\Apollo \parallel b\Apollo
\]

\[
E\text{-CLOSEZAP} \quad F[\text{close } a] \parallel \circ b \parallel a\Apollo \parallel b\Apollo \rightarrow F[\text{raise}] \parallel \circ a \parallel \circ b \parallel a\Apollo \parallel b\Apollo
\]

\[
E\text{-RECEIVEZAP} \quad F[\text{receive } a] \parallel \circ b \parallel a\Apollo \parallel b\Apollo \rightarrow F[\text{raise}] \parallel \circ a \parallel \circ b \parallel a\Apollo \parallel b\Apollo
\]

\[
E\text{-RAISE} \quad F[\text{try } P[\text{raise } \text{ as } x \text{ in } M \text{ otherwise } N]] \rightarrow F[N] \parallel \circ P
\]

\[
E\text{-RAISECHILD} \quad \circ P[\text{raise}] \rightarrow \circ P
\]

\[
E\text{-RAISEMAIN} \quad \circ P[\text{raise}] \rightarrow \text{halt} \parallel \circ P
\]

\[
E\text{-LIFTC} \quad G[C] \rightarrow G[D], \quad \text{if } C \rightarrow D
\]

\[
E\text{-LIFTM} \quad \phi M \rightarrow \phi M', \quad \text{if } M \rightarrow^E M'
\]

Fig. 6. Reduction and Equivalence for Terms and Configurations

does not, and must return the unit value. A configuration has at most one main thread. As well as program threads, configurations include three special forms of thread. A zapper thread (\circ a) manages an endpoint a that has been cancelled, and is used to propagate failure. A halted thread (\text{halt}) arises when the main thread has crashed due to an uncaught exception. A buffer thread (a\Apollo \parallel b\Apollo) models asynchrony: \Apollo and \overrightarrow{\text{W}} are sequences of values ready to be received along endpoints a and b respectively. We sometimes find it useful to distinguish top-level threads \mathcal{T} (main threads and halted threads) from auxiliary threads \mathcal{A} (child threads, zapper threads, and buffer threads).

\textbf{Environments.} We extend type environments \Gamma to include runtime names of session type and introduce runtime type environments \Delta, which type both buffer endpoints of session type and channels of type S^\delta for some S, but not object variables.

\textbf{Contexts.} Evaluation contexts \mathcal{E} are set up for standard left-to-right call-by-value evaluation. Pure contexts \mathcal{P} are those evaluation contexts that include no exception handling frames. Thread
contexts $F$ support reduction in program threads. Configuration contexts $G$ support reduction under $\nu$-binders and parallel composition.

*Free Names.* We let the meta operation $\text{fn}(\cdot)$ denote the set of free names in a term, type environment, buffer environment, value, configuration, pure context, or evaluation context.

*Syntactic Sugar.* We follow the standard convention that parallel composition of configurations associates to the right. We write $\parallel V$, $\parallel P$, and $\parallel E$, as shorthand for the parallel composition of zapper threads for each free name in values $V$, pure contexts $P$, and evaluation contexts $E$, respectively.

Following prior work on linear functional languages with session types [Gay and Vasconcelos 2010; Lindley and Morris 2015, 2016, 2017], we present the semantics of EGV via a deterministic reduction relation on terms ($\beta$-reduction), an equivalence relation on configurations (equiv), and a nondeterministic reduction relation on configurations ($\Rightarrow$). We write $\equiv\Rightarrow$ for the relation $\equiv\beta\equiv$. Fig. 6 presents reduction and equivalence rules for terms and configurations.

**Term Reduction.** Reduction on terms is standard call-by-value $\beta$-reduction.

**Configuration Equivalence.** A running program can make use of the standard structural $\pi$-calculus equivalence rules [Milner 1999] of associativity and commutativity of parallel composition, name restriction reordering, and scope extrusion. Formally, equivalence is defined as the smallest congruence relation satisfying the equivalence axioms in Figure 6. We incorporate a further rule to allow buffers to be treated symmetrically and two garbage collection rules, allowing completed child threads and cancelled empty buffers to be discarded.

**Communication and Concurrency.** The E-Fork rule creates two fresh names for each endpoint of a channel, returning one name and substituting the other in the body of the spawned thread, as well as creating a channel with two empty buffers. The E-Send and E-Receive rules send to and receive from a buffer. The E-Close rule discards an empty buffer once a session is complete.

**Cancellation.** The E-Cancel rule cancels an endpoint by creating a zapper thread. The E-Zap rule ensures that when an endpoint is cancelled, all endpoints in the buffer of the cancelled endpoint are also cancelled: it dequeues a value from the head of the buffer and cancels any endpoints contained within the dequeued value ($\parallel U$). It is applied repeatedly until the buffer is empty.

**Raising Exceptions.** Following Mostrous and Vasconcelos [2014], an exception is raised when it would be otherwise impossible for a communication action to succeed. The E-ReceiveZap rule raises an exception if an attempt is made to receive along an endpoint whose buffer is empty and whose peer endpoint has been cancelled. Similarly, E-CloseZap raises an exception if an attempt is made to close a channel where the peer endpoint has been cancelled. There is no rule for the case where a thread tries to send a value along a cancelled endpoint; the free names in the communicated value must eventually be cancelled, but this is achieved through E-Zap. We choose not to raise an exception in this case since to do so would violate confluence, which we discuss in more detail in §3.4. Not raising exceptions on message sends to dead peers is standard behaviour for languages such as Erlang.

**Handling Exceptions.** The E-Raise rule invokes the otherwise clause if an exception is raised, while also cancelling all endpoints in the enclosing pure context. If an unhandled exception occurs in a child thread, then all free endpoints in the evaluation context are cancelled and the thread is terminated (E-RaiseChild). If the exception is in the main thread then all free endpoints are cancelled and the main thread reduces to halt (E-RaiseMain).
2.4 Synchrony
As we are interested in writing distributed applications, we consider asynchronous session types. However, our semantics adapts straightforwardly to the synchronous setting, where a send to a cancelled peer must also raise an exception:

\[
\Gamma \vdash \text{E-SyncComm } F[[\text{send } V a]] \parallel F'[\text{receive } a] \rightarrow F[a] \parallel F'[(V, a)]
\]

\[
\Gamma \vdash \text{E-SyncSendZAP } F[[\text{send } V a]] \parallel \xi a \rightarrow F[\text{raise}] \parallel \xi V \parallel \xi a \parallel \xi a
\]

\[
\Gamma \vdash \text{E-SyncRecvZAP } F[\text{receive } a] \parallel \xi a \rightarrow F[\text{raise}] \parallel \xi a \parallel \xi a
\]

\[(\nu a)(\xi a) \parallel \xi a \parallel C \equiv C\]

3 METATHEORY
Even in the presence of channel cancellation and exceptions, EGV retains GV’s strong metatheory [Lindley and Morris 2015]. The central property of session-typed systems is session fidelity: all communication follows the prescribed session types. Session fidelity follows as a corollary of preservation of configuration typing under reduction.

Session calculi with roots in linear logic are deadlock-free as interpreting the logical cut rule as a combination of name restriction and parallel composition necessarily ensures acyclicity [Caires and Pfenning 2010]. It is also possible to use deadlock-freedom to derive a global progress result. We prove that global progress holds even in the presence of channel cancellation. (Our proof is direct, not requiring catalyster processes [Carbone et al. 2014; Mostrous and Vasconcelos 2014].) We also prove that EGV is confluent and terminating.

3.1 Runtime Typing
To state our main results we require typing rules for names and configurations. These are given in Fig. 7. The configuration typing judgement has the shape \( \Gamma; \Delta \vdash \phi \ C \), which states that under type environment \( \Gamma \), runtime environment \( \Delta \), and thread flag \( \phi \), configuration \( C \) is well-typed. We additionally require that \( \text{fn}(\Gamma) \cap \text{fn}(\Delta) = \emptyset \). Thread flags ensure that there can be at most one top-level thread which can return a value: \( \bullet \) denotes a configuration with a top-level thread and \( \circ \) denotes a configuration without. The main thread returns the result of running a program. Any configuration \( C \) such that \( \Gamma; \Delta \vdash \bullet \ C \) has exactly one main thread or halted thread as a subconfiguration. We write \( \Gamma; \Delta \vdash \bullet \ C : A \) whenever the derivation of \( \Gamma; \Delta \vdash \bullet \ C \) contains a subderivation of the form

\[
\frac{\Gamma' \vdash M : A}{\Gamma'; \Phi \bullet M \quad \text{or} \quad \vdash \Phi \text{-halt}}
\]

We say that a \( C \) is a ground configuration if there exists \( A \) such that \( \vdash \Phi \ C : A \) and \( A \) contains no session types or function types.

The T-Nu rule introduces a channel name; T-CONNECT1 and T-CONNECT2 connect two configurations over a channel; and T-MIX composes two configurations that share no channels. The latter three rules use the + operator to combine the flags from subconfigurations. The T-MAIN and T-CHILD rules introduce main and child threads. Child threads always return the unit value. The T-HALT rule types the halt configuration, which signifies that an unhandled exception has occurred in the main thread. The T-ZAP rule types a zapper thread, given a single name in the type environment. The T-BUFFER rule ensures that buffers contain values corresponding to the session types of their endpoints. This is the only rule that consumes names from the runtime environment. Buffers rely on two auxiliary judgements. The queue typing judgement \( \Gamma \vdash \nu \nu : \nu A \) states that under type environment \( \Gamma \), the sequence of values \( \nu \nu \) have types \( \nu A \). The session slicing operator \( S/A \) captures reasoning about session types discounting values contained in the buffer. The session
The partiality of the slicing operator coupled with the duality constraint ensures that at least one subconfiguration typeability and replacement. The proof cases can be found in Appendix C.1. □

Preservation for the functional fragment of EGV is standard.

**Lemma 3.1 (Preservation (Terms)).** If \( \Gamma \vdash M : A \) and \( M \xrightarrow{M} M' \), then \( \Gamma \vdash M' : A \).

**Theorem 3.2 (Preservation).** If \( \Psi; \Delta \vdash \phi C \) and \( C \xrightarrow{C} C' \), then there exist \( \Psi', \Delta' \) such that \( \Psi; \Delta \xrightarrow{\phi} \Psi'; \Delta' \) and \( \Psi' \vdash \phi C' \).

**Proof.** By induction on the derivation of \( C \xrightarrow{C} C' \), making use of Lemma 3.1, and lemmas for subconfiguration typeability and replacement. The proof cases can be found in Appendix C.1.

---

**Fig. 7. Runtime Typing**

<table>
<thead>
<tr>
<th>Term Typing</th>
<th>Session Slicing</th>
<th>Queue Typing</th>
<th>Configuration Typing</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Gamma \vdash M : A )</td>
<td>( S/\epsilon = S )</td>
<td>( \Gamma \vdash \overline{V} : \overline{A} )</td>
<td>( \Gamma; \Delta \vdash \phi C )</td>
</tr>
<tr>
<td><strong>T-Name</strong></td>
<td>( a : S \vdash a : S )</td>
<td>( + \epsilon : \epsilon )</td>
<td><strong>T-Nu</strong></td>
</tr>
<tr>
<td>( \Gamma; \Delta, a : S^# \vdash \phi C )</td>
<td>( \Gamma; \Delta \vdash (\nu a)C )</td>
<td>( \Gamma_1; \Gamma_2; \Delta_1, \Delta_2, a : S^# \vdash \phi_1 + \phi_2 C \</td>
<td></td>
</tr>
<tr>
<td>( \Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^# \vdash \phi_1 + \phi_2 C \</td>
<td></td>
<td>D )</td>
<td></td>
</tr>
</tbody>
</table>

**Preservation**

**3.2 Preservation**

The partiality of the slicing operator coupled with the duality constraint ensures that at least one queue in a buffer is always empty. Appendix A shows an example configuration typing derivation.
Typing and Configuration Equivalence. As is common in logically-inspired session-typed functional languages [Lindley and Morris 2015, 2017], typeability of configurations is not preserved by equivalence. Consider $\Gamma; \Delta \vdash^\phi (va)(vb)(C || (D || E))$ with $a \in \text{fn}(C), b \in \text{fn}(D)$, and $a, b \in \text{fn}(E)$. But $\Gamma; \Delta \not\vdash^\phi (va)(vb)((C || D) || E)$. Fortunately this looseness of the equivalence relation is unproblematic: we may always safely re-associate parallel composition (for example, $\Gamma; \Delta \vdash^\phi (va)(vb)((C || E) || D)$; see Appendix C.1), and any reduction sequence which uses ill-typed equivalences may be replaced by one that does not.

**Theorem 3.3 (Preservation Modulo Equivalence).** If $\Psi; \Delta \vdash^\phi C, C \equiv D$, and $D \longrightarrow D'$, then:

1. There exists some $E \equiv D$ and some $E'$ such that $\Psi; \Delta \vdash^\phi E$ and $E \longrightarrow E'$
2. There exist $\Psi', \Delta'$ such that $\Psi; \Delta \longrightarrow^\psi \Psi'$; $\Delta'$ and $\Psi'$; $\Delta' \vdash^\phi E'$
3. $D' \equiv E'$

**Proof.** The only non-trivial reductions are those involving a synchronisation with a buffer (E-Send, E-Receive, E-Close, E-Zap, E-CloseZap, E-ReceiveZap). The only equivalence rule that can lead to an ill-typed configuration is associativity of parallel composition $C || (D || E) \equiv (C || D) || E$ where both compositions arise from the T-Connect$_1$ and T-Connect$_2$ rules. The only reason to apply the associativity rule from left-to-right is to enable threads inside $C$ and $D$ to synchronise. But for synchronisation to be possible there must exist a name $a$ such that $a \in \text{fn}(C)$ and $a \in \text{fn}(D)$. Because the left-hand-side of the equation is well-typed, we know that $C$ and $E$ have no names in common, that $D$ and $E$ share a name, and that the right-hand-side must be well-typed as there is still exactly one channel connecting each of the parallel compositions. The argument for applying the rule from right-to-left is symmetric. In summary, any ill-typed use of equivalence is useless. □

### 3.3 Progress

To prove that EGV enjoys a strong notion of progress we identify a canonical form for configurations. We prove that every well-typed configuration is equivalent to a well-typed configuration in canonical form, and that ground configurations can always either reduce, or are equivalent to either a value or halt.

The functional fragment of EGV enjoys progress.

**Lemma 3.4 (Progress: Open Terms).** If $\Psi \vdash M : A$, then either:

- $M$ is a value;
- there exists some $M'$ such that $M \longrightarrow_M M'$; or
- $M$ has the form $E[M']$, where $M'$ is a session typing primitive of the form: fork $V$, send $V$ $W$, receive $V$, close $V$, or cancel $V$.

**Proof.** By induction on the derivation of $\Psi \vdash M : A$. □

To reason about progress of configurations, we characterise canonical forms, which make explicit the property that at most one name is shared between threads. Recall that $\mathcal{A}$ ranges over auxiliary threads and $\mathcal{T}$ over top-level threads (Fig. 5). Let $\mathcal{M}$ range over configurations of the form:

$$\mathcal{A}_1 || \cdots || \mathcal{A}_n || \mathcal{T}$$

**Definition 3.5 (Canonical Form).** A configuration $C$ is in canonical form if there is a sequence of names $a_1, \ldots, a_n$, a sequence of configurations $\mathcal{A}_1, \ldots, \mathcal{A}_n$, and a configuration $M$, such that:

$$C = (va_1)(\mathcal{A}_1 || (va_2)(\mathcal{A}_2 || \cdots || (va_n)(\mathcal{A}_n || M) \ldots))$$

where $a_i \in \text{fn}(\mathcal{A}_i)$ for each $i \in 1..n$.}

The following lemma implies that communication topologies are always acyclic.

**Lemma 3.6.** If $\Gamma; \Delta \vdash^\Psi C$ and $C = G[D \parallel E]$, then $\text{fn}(D) \cap \text{fn}(E)$ is either $\emptyset$ or $\{a\}$ for some $a$.

**Proof.** By induction on the derivation of $\Gamma; \Delta \vdash^\Psi C$; the only interesting rules are those for parallel composition. As the environments are well-formed, $\text{fn}(\Gamma) \cap \text{fn}(\Delta) = \emptyset$. Thus, $T$-Connect$_1$ and $T$-Connect$_2$ allow exactly one name to be shared, whereas $T$-Mix forbids sharing of names. $\square$

All well-typed configurations can be written in canonical form.

**Theorem 3.7 (Canonical Forms).** Given $C$ such that $\Gamma; \Delta \vdash^\Psi C$, there exists some $D \equiv C$ such that $\Gamma; \Delta \vdash^\Psi D$ and $D$ is in canonical form.

**Proof.** By induction on the count of $\nu$-bound variables, following Lindley and Morris [2015] and making use of Lemma 3.6. The additional features of EGV do not change the essential argument. The full proof can be found in Appendix C.2. $\square$

**Definition 3.8.** We say that term $M$ is ready to perform an action on name $a$ if $M$ is about to send on, receive on, close, or cancel $a$. Formally:

$$\text{ready}(a, M) \triangleq \exists \nu. (M = E[send\ V\ a]) \lor (M = E[receive\ a]) \lor (M = E[close\ a]) \lor (M = E[cancel\ a])$$

Using the notion of a ready thread, we may classify a notion of progress for open configurations.

**Theorem 3.9 (Progress: Open).** Suppose $\Psi; \Delta \vdash^\cdot C$, where $C$ is in canonical form.

Let $C = (va_1)(A_1 \parallel (va_2)(A_2 \parallel \cdots \parallel (va_n)(A_n \parallel M) \cdots))$.

Either there exists some $C'$ such that $C \Rightarrow C'$, or:

1. For $1 \leq i \leq n$, each auxiliary thread $A_i$ is either:
   1. A child thread $\circ M$ for which there exists $a \in \{a_j \mid 1 \leq j \leq i\} \cup \text{fn}(\Psi)$ such that $\text{ready}(a, M)$;
   2. A zapper thread $\OPT a$; or
   3. A buffer.

2. $M = A' \parallel \cdots \parallel A'' \parallel T$ such that for $1 \leq j \leq m$:
   1. $A'_j$ is either:
      1. A child thread $\circ N$ with $N = ()$ or $\text{ready}(a, N)$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$;
      2. A zapper thread $\OPT a$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$; or
      3. A buffer.
   2. Either $T = \OPT N$, where $N$ is either a value or $\text{ready}(a, N)$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$; or $T = \text{halt}$.

**Proof.** The result follows from a more verbose, but finer-grained, property which we prove by induction on the derivation of $\Psi; \Delta \vdash^\cdot C$. Full details are in Appendix C.3. $\square$

This theorem tells us that open reduction cannot "go wrong". A progress theorem states that either reduction is possible or the configuration is a value. Conditions 1(a)(b)(c) and 2(a)(b) constitute a suitable generalisation of "value".

By restricting attention to closed environments, we obtain a tighter progress property.

**Theorem 3.10 (Progress: Closed).** Suppose $\cdot \vdash^\cdot C$ where $C$ is in canonical form.

Let $C = (va_1)(A_1 \parallel (va_2)(A_2 \parallel \cdots \parallel (va_n)(A_n \parallel M) \cdots))$.

Either there exists some $C'$ such that $C \Rightarrow C'$, or:

1. For $1 \leq i \leq n$, each auxiliary thread $A_i$ is either:
   1. A child thread $\circ M$ for some $M$ such that $\text{ready}(a_i, M)$; or
   2. A zapper thread $\OPT a_i$; or
   3. A buffer.
(2) Either $M = \bullet W$ for some value $W$, or $M = \text{halt}$.

The above progress results do not specifically mention deadlock. However, Lemma 3.6 ensures deadlock-freedom. Nevertheless, communication can still be blocked if an endpoint appears in the value returned by the main thread. A conservative way of disallowing endpoints in the result is to insist that the return type of the program be free of session types and function types (closures may capture endpoints). All configurations of such a programs are ground configurations.

**Theorem 3.11 (Global Progress).** Suppose $C$ is a ground configuration. Either there exists some $C'$ such that $C \rightarrow C'$; or $C \equiv \bullet V$; or $C \equiv \text{halt}$.

**Proof.** As a consequence of Theorem 3.10, either there exists some $C'$ such that $C \rightarrow C'$, or $C \Rightarrow$ and each thread $A_i$ must be a zapper, a buffer, or ready to perform an action. If $C \Rightarrow$, since $C$ is ground, by Lemma 3.6, we have that no thread can be ready to perform an action. Thus, each $A_i$ must be either $\epsilon()$, a zapper, or an empty buffer. The result then follows by the garbage collection congruences of Fig. 6. □

### 3.4 Confluence

EGV enjoys a strong form of confluence known as the diamond property [Barendregt 1984].

**Theorem 3.12 (Diamond Property).** If $\Psi; \Delta \vdash^\phi C$, and $C \Rightarrow D_1$, and $C \Rightarrow D_2$, then either $D_1 \equiv D_2$, or there exists some $D_3$ such that $D_2 \Rightarrow D_3$ and $D_2 \Rightarrow D_3$.

**Proof.** First, note that $\Rightarrow^* M$ is entirely deterministic and hence confluent due to the call-by-value, left-to-right ordering imposed by evaluation contexts. By linearity, we know that endpoints to different buffers may not be shared, so it follows that communication actions on different channels may be performed in any order. Asynchrony and cancellation introduce two critical pairs which may be resolved in a single step; see Appendix C.4 for details.

**Remark.** The system becomes non-confluent if we choose to raise an exception when sending to a cancelled buffer. Suppose that instead of the current semantics, we were to replace E-Send with the following two rules:

\[
(\nu b)(F[send\ U\ a] \parallel a(\nu b) \rightarrow b(W)\parallel \phi M) \rightarrow (\nu b)(F[a] \parallel a(\nu b) \rightarrow b(W) \cdot U)\parallel \phi M)
\]

\[
F[send\ U\ a] \parallel \epsilon b \parallel a(\nu b) \rightarrow b(W) \rightarrow F[raise] \parallel \epsilon b \parallel \epsilon U \parallel a(\nu b) \rightarrow b(W)
\]

Then, sending and cancelling peer endpoints of a buffer results in a non-convergent critical pair:

\[
(\nu b)(F[send\ U\ a] \parallel F'[cancel\ b] \parallel a(V) \rightarrow b(W))
\]

\[
(\nu b)(F[a] \parallel F'[cancel\ b] \parallel a(V) \rightarrow b(W) \cdot U)
\]

\[
(\nu b)(F[a] \parallel F'[cancel\ b] \parallel a(V) \rightarrow b(W))
\]

\[
(\nu b)(F[send\ U\ a] \parallel F'[raise] \parallel F'[cancel\ b] \parallel a(V) \rightarrow b(W))
\]

\[
(\nu b)(F[send\ U\ a] \parallel F'[raise] \parallel a(V) \rightarrow b(W) \cdot U)
\]

In either case, the endpoints contained in $U$ will still eventually be cancelled, thus preservation and global progress still hold. However, the lack of confluence affects exactly when the exception is raised in context $F$. This decision has practical significance, in that it characterises the race between sending a message and propagating a cancellation notification.

### 3.5 Termination

As EGV is linear, it has an elementary strong normalisation proof.

**Theorem 3.13 (Strong Normalisation).** If $\Psi; \Delta \vdash^\phi C$, then there are no infinite $\Rightarrow$ reduction sequences from $C$.  

---

We believe that a logical relations argument along the lines of Pérez et al. [2012] or a CPS translation continues to hold in the presence of unrestricted types or shared channels for session initiation, but the proof technique is necessarily more involved. We believe that a logical relations argument along the lines of Lindley and Morris [2016] would suffice.

4 EXTENSIONS

4.1 User-defined Exceptions with Payloads

In order to focus on the interplay between exceptions and session types we have thus far considered handling a single kind of exception. In practice it can be useful to distinguish between multiple kinds of user-defined exception, each of which may carry a payload.

Consider again handling the exception in checkDetails. An exception may arise if the database is corrupt, or if there are too many connections. We might like to handle each case separately:

\[
\text{exnServer4}(s) \triangleq \\
\quad \text{let } ((\text{username}, \text{password}), s) = \text{receive } s \text{ in} \\
\quad \text{try checkDetails(\text{username}, \text{password}) as res in} \\
\quad \quad \text{if res then let } s = \text{select } \text{Authenticated } s \text{ in serverBody}(s) \\
\quad \quad \text{else let } s = \text{select } \text{AccessDenied } s \text{ in close } s \\
\quad \text{unless} \\
\quad \text{DBCorrupt}(y) \quad \Rightarrow \text{cancel } s; \log("\text{Database Corrupt: }" + y) \\
\quad \text{TooManyConnections}(y) \Rightarrow \text{cancel } s; \log("\text{Too many connections: }" + y)
\]

An exception in checkDetails might be raised by the term \text{raise} DatabaseCorrupt(filename), for example. Our approach generalises straightforwardly to handle this example.

Syntax. Figure 8 shows extensions to EGV for exceptions with payloads. We introduce a type of exceptions, Exn. We assume a countably infinite set \(X \in \mathbb{E}\) of exception names, and a type schema function \(\Sigma(X) = A\) mapping exception names to payload types. We extend \text{raise} to take a term of type Exn as its argument. Finally, we generalise \text{try} \text{Lax} in \text{Otherwise} \ N to \text{try} \text{Lax} in \text{M} \text{unless} \ H, where \(H\) is an exception handler with clauses \(\{X_i(y_i) \mapsto N_i\}_i\), such that \(X_i\) is an exception name; \(y_i\) binds the payload; and \(N_i\) is the clause to be evaluated when the exception is raised.

Typing Rules. The TP-EXN rule ensures that an exception’s payload matches its expected type. The TP-RAISE and TP-TRY are the general extensions of T-RAISE and T-TRY.

Semantics. Our presentation is similar to operational accounts of effect handlers; the formulation here is inspired by that of Hillerström et al. [2017]. To define the semantics of the generalised exception handling construct, we first introduce the auxiliary function \text{handled}(E), which defines the exceptions handled in a given evaluation context:

\[
\begin{align*}
\text{handled}(P) &= \emptyset \\
\text{handled}(\text{try } E \text{ as } x \text{ in } M \text{ unless } H) &= \text{handled}(E) \cup \text{dom}(H) \\
\text{handled}(E) &= \text{handled}(E'), \quad \text{if } E \text{ is not a } \text{try} \text{ and } E' \text{ is the immediate subcontext of } E
\end{align*}
\]

The EP-RAISE rule handles an exception. The side conditions ensure that the exception is caught by the nearest matching handler and is handled by the appropriate clause. As with plain EGV, all free names are safely discarded. The EP-RAISECHILD and EP-RAISEMAIN rules cover the cases where an exception is unhandled. Due to the use of the handled function we no longer require pure contexts.
Syntax

Types

\[ A, B ::= \cdots \mid \text{Exn} \]

Terms

\[ L, M, N ::= \cdots \mid X(M) \mid \text{raise } M \mid \text{try } L \text{ as } x \text{ in } M \text{ unless } H \]

Exception Handlers

\[ H ::= \{ X_i(x_i) \mapsto N_i \}_i \]

Runtime Syntax

Evaluation Contexts

\[ E ::= \cdots \mid \text{raise } E \mid \text{try } E \text{ as } x \text{ in } M \text{ unless } H \]

Term typing

\[ \Sigma(X) = A \quad \Gamma \vdash M : A \]

\[ \Gamma \vdash X(M) : \text{Exn} \quad \Gamma \vdash \text{raise } M : A \]

\[ \Gamma \vdash \text{raise } L \text{ as } x \text{ in } M \text{ unless } \{ X_i(y_i) \mapsto N_i \}_i : B \]

Term and Configuration Reduction

\[ M \rightarrow_M N \quad C \rightarrow_D \]

All of EGV’s metatheoretic properties (preservation, global progress, confluence, and termination) adapt straightforwardly to this extension.

4.2 Unrestricted Types and Access Points

Unrestricted (intuitionistic) types allow some values to be used in a non-linear fashion. Access points [Gay and Vasconcelos 2010] provide a more flexible method of session initiation than fork, allowing two threads to dynamically establish a session. Both features are useful in practice: unrestricted types because some data is naturally multi-use, and access points because they admit cyclic communication topologies supporting racey stateful servers such as chat servers. Access points decouple spawning a thread from establishing a session. An access point has the unrestricted type AP(S); we write un(A) to mean that A is unrestricted and un(\(\Gamma\)) if un(A\(_i\)) for all \(x_i : A_i \in \Gamma\).

Figure 9 shows the syntax, typing rules, and reduction rules for EGV extended with access points.

\textit{Unrestricted Types.} To support unrestricted types, we introduce a splitting judgement (\(\Gamma = \Gamma_1 + \Gamma_2\)), which allows variables of unrestricted type to be shared across sub-environments, but requires linear variables to be used only in a single sub-environment. We relax rule T-Var to allow the use of unrestricted environments, and adapt all rules containing multiple subterms to use the splitting judgement. We detail T-App in the figure; the adaptations of other rules are similar. While unrestricted types are useful in general, we show the specific case of unrestricted access points.

\textit{Access points.} The \texttt{spawn } M \text{ construct spawns } M \text{ as a new thread, \texttt{newS} }\text{ creates a fresh access point, and \texttt{request } M \text{ and \texttt{accept } M} \text{ generate fresh endpoints that are matched up nondeterministically to form channels. With access points we can macro-express fork:}

\[ \text{fork } M \triangleq \text{let } ap = \text{newS in spawn } (M \text{ (accept ap)}); \text{ request ap} \]
Fig. 9. Access Points

Reduction rules. We let $z$ range over access point names. Configuration $(vz)C$ denotes binding access point name $z$ in $C$, and $z(X, Y)$ is an access point with name $z$ and two sets $X$ and $Y$ containing endpoints to be matched.

Rule E-Spawn creates a new child thread but, unlike fork, returns the unit value instead of creating a channel and returning an endpoint. Rule E-New creates a new access point with fresh name $z$. Rules E-Accept and E-Request create a fresh name $a$, returning the newly-created name to the thread, and adding the name to sets $X$ and $Y$ respectively. Rule E-Match matches two endpoints $a$ and $b$ contained in $X$ and $Y$, and creates an empty buffer $a(e) \cdots b(e)$.

Configuration typing. Configuration typing judgements again have the shape $\Gamma; \Delta + \phi C$. Whereas $\Gamma$ may contain unrestricted variables, $\Delta$ remains entirely linear.
Read bottom-up, rule TA-ApName adds an unrestricted reference $z : \text{AP}(S)$ to $\Gamma$, and a linear entry $z : S$ to $\Delta$. Rule TA-Ap types an access point configuration. We write $X : S$ for $a_1 : S, \ldots, a_n : S$, where $X = \{a_1, \ldots, a_n\}$. For an access point $z(X, Y)$ to be well-typed, $\Delta$ must contain $z : S$, along with the names in $X$ having type $\text{AP}(S)$ and the names in $Y$ having type $S$. Rule T-Connect generalises T-Connect$_1$ and T-Connect$_2$ to allow any number of channels to communicate across a buffer; this therefore introduces the possibility of deadlock.

**Interaction with cancellation.** We need no additional reduction rules to account for interaction between access points and channel cancellation. Should an endpoint waiting to be matched be cancelled, it is paired as usual, and interaction with its associated buffer raises an exception:

$$\frac{\not\epsilon a \parallel F[\text{receive } b] \parallel z(\{a\}, \{b\}) \Rightarrow \not\epsilon a \parallel F[\text{receive } b] \parallel z(\epsilon, \epsilon) \parallel a(\epsilon) \sim b(\epsilon)}{\not\epsilon a \parallel F[\text{raise}] \parallel \not\epsilon b \parallel z(\epsilon, \epsilon) \parallel a(\epsilon) \sim b(\epsilon)}$$

**Metatheory.** By decoupling process and channel creation we lose the guarantee that the communication topology is acyclic, and therefore introduce the possibility of deadlock. Preservation continues to hold—in fact, we gain a stronger preservation result since the use of TA-ConnectN allows typeability to be preserved by equivalences.

**Theorem 4.1 (Preservation Modulo Equivalence (Access Points)).**

If $\Psi; \Delta \vdash^\phi C$ and $C \Rightarrow D$, then there exist $\Psi'; \Delta'$ such that $\Psi; \Delta \Rightarrow \Psi'; \Delta'$ and $\Psi'; \Delta' \vdash^\phi D$.

**Proof.** By induction on the derivation of $C \Rightarrow D$ and preservation by $\equiv$; see Appendix D. $\square$

Alas, the introduction of cyclic topologies and therefore the loss of deadlock-freedom necessarily violates global progress. Nevertheless, a weaker form of progress still holds: if a configuration does not reduce, then it is due to deadlock rather than cancellation.

**Theorem 4.2 (Progress (Access Points)).** Suppose $\vdash^\phi C$ and $C \Rightarrow$. Then each thread in $C$ is either a value; a buffer; a zapper thread; an access point; requesting or accepting on an access point; or ready to perform a communication action.

If $C$ contains a thread $\phi M$ and ready$(a, M)$ for some name $a$, then $C$ contains some buffer $a(\epsilon) \sim b(\overrightarrow{\mathcal{W}})$, and $C$ does not contain a zapper thread $\not\epsilon b$.

**Proof.** We can prove a similar property for open configurations by induction on the derivation of $\Psi; \Delta \vdash^\phi C$; the above result arises as a corollary and by inspection of the reduction rules. $\square$

In the presence of access points confluence and termination no longer hold: access points are nondeterministic and can encode higher-order state and hence fixpoints via Landin’s knot.

### 4.3 Recursive Session Types

Recursive session types support repeating protocols. The extension of EGV with recursive session types is standard [Lindley and Morris 2016, 2017] and orthogonal to the main ideas of this paper, so we do not spell out the details here. The implementation (§5) does provide recursive session types.

### 5 SESSION TYPES WITHOUT TIERS

In this section we describe our extensions to Links to support exception handling, as well as extensions to the Links concurrency runtimes to support distribution. Links [Cooper et al. 2007] is a statically-typed, ML-inspired, impure functional programming language designed for the web. Links is designed to allow code for all "tiers" of a web application—client, server, and database—to be written in a single language. Lindley and Morris [2017] extend Links with first-class session types, relying on lightweight linear typing [Mazurak et al. 2010] and row polymorphism [Rémy 2007] to allow any number of channels to communicate across a buffer; this therefore introduces the possibility of deadlock.
We extend their work to account for distributed web applications, which amongst other things necessitates handling failure.

5.1 The Links Model

Links provides a uniform language for web applications. Client code is compiled to JavaScript, server code is interpreted, and database queries are compiled to SQL. Each client and server has its own concurrency runtime, providing lightweight processes and message passing communication. Earlier versions of Links [Cooper et al. 2007] invoked a fresh copy of the server per server request and communication between client and server was via RPC calls. Advances such as WebSockets allow richer applications where data (for example, comments on a GitHub pull request) flows more freely between client and server. Moving to a model based on lightweight threads and session-typed channels avoids the inversion of control inherent in RPC-style systems, and allows development to be driven by the communication protocol.

Links now adopts a persistent application server model, incorporating client-server communication using session-typed channels. Since channels are a location-transparent abstraction, we also optionally allow the abstraction of client-to-client communication, routed through the server.

5.2 Concurrency

Links provides typed actor-style concurrency where processes have a single incoming message queue and can send asynchronous messages. Lindley and Morris [2017] extend Links with session-typed channels, using Links’ process-based model but replacing actor mailboxes with session-typed channels. We extend their implementation to support distribution and failure handling.

The client relies on continuation-passing style (CPS), trampolining, and co-operative threading. Client code is compiled to CPS, and explicit yield instructions are inserted at every function application. When a process has yielded a given number of times, the continuation is pushed to the back of a queue, and the next process is pulled from the front of the queue. While modern browsers are beginning to integrate tail-recursion, and we have updated the Links library to support it, adoption is not yet widespread. Thus, we periodically discard the call stack using a trampoline. Cooper [2009] discusses the Links client concurrency model in depth. The server implements concurrency on top of the OCaml lwt library [Vouillon 2008], which provides lightweight co-operative threading. At runtime, a channel is represented as a pair of endpoint identifiers:

\[(\text{Peer endpoint}, \text{Local endpoint})\]

Endpoint identifiers are unique. If a channel \((a, b)\) exists at a given location, then that location should contain a buffer for \(b\).

5.3 Distributed Communication

To support bidirectional communication between client and server we use WebSockets [Fette and Melnikov 2011]. A WebSocket connection is established by a client. When a request is made and a web page is generated, each client is assigned a unique identifier, which it uses to establish a WebSocket connection. Any messages the server attempts to send prior to a WebSocket connection being established are buffered and delivered after the connection is established. We use a JSON protocol to communicate messages such as access point operations, remote session messages, and endpoint cancellation notifications.

It is possible that one client will hold one endpoint of a channel, and another client will hold the other endpoint. In order to provide the illusion of client-to-client communication, we route the
communication between the two clients via the server. The server maintains a map

\[
\text{Endpoint ID} \mapsto \text{Location}
\]

where Location is either Server or Client(ID), where ID identifies a particular client. The map is updated if a new connection is established; an endpoint is sent as part of a message; or a client disconnects. The server also maintains a map

\[
\text{Client ID} \mapsto [\text{Channel}]
\]

associating each client with the publicly-facing channels residing on that client, where Channel is a pair of endpoints \((a, b)\) such that \(b\) is the endpoint residing on the client. Much like TCP connections, WebSocket connections raise an event when a connection is disconnected. Upon receiving such an event, all channels associated with the client are cancelled, and exceptions are invoked as per the exception handling mechanism described in §2 and §5.4.

**Distributed Delegation.** It is possible to send endpoints as part of a message. Session delegation in the presence of distributed communication requires some care to ensure that messages are delivered to the correct participant; our implementation adapts the algorithms of Hu et al. [2008]. Further details can be found in Appendix E.

### 5.4 Session Typing with Failure Handling

**Effect Handlers.** Effect handlers [Plotkin and Pretnar 2013] provide a modular approach to programming with user-defined effects. Exception handlers are a special case of effect handlers. Consequently, we leverage the existing implementation of effect handlers in Links [Hillerström and Lindley 2016; Hillerström et al. 2017]. In §4 we generalise try – as – in – otherwise– to accommodate user defined exceptions. Effect handlers generalise further to support what amounts to resumable exceptions in which the handler has access not only to a payload, but also the delimited continuation (i.e. evaluation context) from the point at which the exception was raised up to the handler, allowing effect handlers to implement arbitrary side-effects; not just exceptions. We translate exception handling as follows.

\[
\begin{align*}
\text{[raise]} &= \text{do raise} \\
\text{[try } L \text{ as } x \text{ in } M \text{ otherwise } N] &= \text{handle } [L] \text{ with} \\
&\quad \text{return } x \mapsto [M] \\
&\quad \text{raise } r \mapsto \text{cancel } r; [N]
\end{align*}
\]

The introduction form \textbf{do} \textbf{op} invokes an operation \textbf{op} (which may represent raising an exception or some other effect). The elimination form \textbf{handle} \textbf{M} \textbf{with} \textbf{H} runs effect handler \textbf{H} on the computation \textbf{M}. In general an effect handler \textbf{H} consists of a \textit{return clause} of the form \textbf{return} \textit{x} \mapsto N, which behaves just like the success continuation \((x \text{ in } N)\) of an exception handler, and a collection of \textit{operation clauses}, each of the form \textbf{op} \textit{p} \textit{r} \mapsto N, specifying how to handle an operation analogously to how exception handler clauses specify how to handle an exception, except that as well as binding payload parameters \textit{p}, an operation clause also binds a \textit{resumption} parameter \textit{r}. The resumption \textit{r} binds a closure representing the continuation up to the nearest enclosing effect handler, allowing control to pass back to the program after handling the effect. In the case of our translation, the raise operation has no payload, and rather than invoking the resumption \textit{r} we cancel it, assuming the natural extension of cancellation to arbitrary linear values, whereby all free names in the value are cancelled (\textit{r} being bound to the current evaluation context reified as a value). A formalisation of linear effect handlers for session typing is outside the scope of this paper and left as future work.
Raising exceptions. An exception may be raised either explicitly through an invocation of `raise` (desugared to `do raise`), or through a blocked `receive` call where the peer endpoint has been cancelled. Thus, we know statically where any exceptions may be raised. To support cancellation of closures on the client, we adorn closures with an explicit environment field that can be directly inspected. Currently, Links does not closure-convert continuations on the client, so we use a workaround to simulate cancelling a resumption (as required by the translation \([-\)]). When compiling client code, for each occurrence of `do raise`, we compile a function that inspects all affected variables and cancels any affected endpoints in the continuation. For each occurrence of `receive`, we compile a continuation to cancel affected endpoints to be invoked by the runtime system if the receive operation fails.

5.5 Distributed Exceptions

Our implementation fully supports the semantics described in §2. The concurrency runtime at each location maintains a set of cancelled endpoints.

Cancellation. Suppose endpoint \(a\) is connected to peer endpoint \(b\). If \(a\) is cancelled, then all endpoints in the queue for \(a\) are also cancelled according to the E-Zap rule. If \(a\) and \(b\) are at the same location, then \(a\) is added to the set of cancelled endpoints. If they are at different locations, then a cancellation notification for \(a\) is routed to \(b\)'s location. Zapper threads are modelled in the implementation by recording sets of cancelled endpoints and propagating cancellation messages.

Failed communications. Again, suppose endpoint \(a\) is connected to peer endpoint \(b\). Should a process attempt to read from \(a\) when the buffer for \(a\) is empty, then the runtime will check to see whether \(b\) is in the set of cancelled endpoints. If so, then \(a\) is cancelled and an exception is raised in the blocked process; if not, the process is suspended until a message is ready. Should the runtime later add \(b\) to the set of cancelled endpoints, then again \(a\) is cancelled and an exception raised. These actions implement the E-ReceiveZap rule.

Disconnection. To handle disconnection, the server maintains a map from client IDs to the list of endpoints at the associated client. WebSockets—much like TCP sockets—raise a closed event on disconnection. Consequently, when a connection is closed, the runtime looks up the endpoints owned by the terminated client and notifies all other clients containing the peer endpoints.

6 EXAMPLE: A CHAT APPLICATION

In this section we outline the design and implementation of a web-based chat application in Links making use of distributed session-typed channels. We write the following informal specification:

- To initialise, a client must:
  - connect to the chat server; then
  - send a nickname; then
  - receive the current topic and list of nicknames.
- After initialisation the client is connected and can:
  - send a chat message to the room; or
  - change the room’s topic; or
  - receive messages from other users; or
  - receive changes of topic from other users.
- Clients cannot connect with a nickname that is already in use in the room.
- All participants should be notified whenever a participant joins or leaves the room.
Session Types. We can encode much of the specification more precisely as a session type, as shown in Figure 10. The client begins by sending a nickname, and then offers the server a choice of a Join message or a Nope message. In the former case, the client then receives a triple containing the current topic, a list of existing nicknames, and an endpoint (of type ClientReceive) for receiving further updates from the server; and may then continue to send messages to the server as a connected client endpoint (of type ClientSend). (Observe the essential use of session delegation.) In the latter case, communication is terminated. The intention is that the server will respond with Nope if a client with the supplied nickname is already in the chat room (the details of this check are part of the implementation, not part of the communication protocol).

The ClientReceive endpoint allows the client to offer a choice of four different messages: Join, Chat, NewTopic, or Leave. In each case the client then receives a payload (depending on the choice, a nickname, pair of nickname and chat message, or topic change) before offering another choice. The ClientSend endpoint allows the client to select between two different messages: Chat and NewTopic. In each case the client subsequently sends a payload (a chat message or a new topic) before selecting another choice. The chat server communicates with the client along endpoints with dual types.

How can session types help? The connect function (Fig. 11a) is run when a client enters a nickname. First, the client requests a fresh channel of type ChatClient from access point wap of type AP(ChatServer). Next, the client obtains the content of the DOM input box for the nickname by calling getInputContents(nameBoxId), where nameBoxId is the DOM ID for the nickname entry box. Next, the client sends the nickname to the server and waits for a response; in the case of a Join message, the client receives the room data and an incoming message channel, and calls the beginChat function. In the case of a Nope message, an error is printed and the session ends.

Now, suppose the developer forgets to write code to check the server response (Fig. 11b). This implementation is incorrect since there is a communication mismatch: the server is expecting to
fun connect() {
  var s = request(wap);
  var nick = getInputContents(nameBoxId);
  var s = send(nick, s);
  offer(s) {
    case Join(s) ->
      var ((topic, nicks, incoming), s) =
        receive(s);
      beginChat(topic, nicks, incoming, s)
    case Nope(s) ->
      print("Nickname \\
        \"^^ nick \"\" already taken")
  }
}

(a) Correct connect function

(b) Incorrect connect function

Fig. 11. Implementations of connect function

accept or reject the request to join the room, whereas the client is expecting to receive data about
the room. However, since s has type ChatClient but does not follow the protocol, Links catches the
communication mismatch statically. Similarly, Links will statically detect an unused endpoint (e.g.
the developer forgets to finish a protocol) or an endpoint being used more than once, as in §1.2.

Architecture. Figure 12a depicts the architecture of the chat application. Each client has a process
which sends messages over a distributed session channel of type ClientSend to its own worker
process on the server, which in turn sends internal messages to a supervisor process containing the
state of the chat room. These messages trigger the supervisor process to broadcast a message to all
chat clients over a channel of type ~ClientReceive. As is evident from the figure, the communication
topology is cyclic; in order to construct this topology the code makes essential use of access points.

Disconnection. Figure 12b shows the implementation of a worker process which receives messages
from a client. The worker takes the nickname of the client, as well as a channel endpoint of type
WorkerReceive (which is the dual of ClientSend). The server offers the client a choice of sending a
message (Chat), or changing topic (NewTopic); in each case, the associated data is received and a
message dispatched to the supervisor process by calling chat or newTopic. When a client closes its
count to the server, all associated endpoints are cancelled. Consequently, an exception will
be raised when evaluating the offer or receive expressions. To handle disconnection, we wrap the
function in an exception handler, which recursively calls worker if the interaction is successful, and
notifies the supervisor that the user has left via a call to leave if an exception is raised.

Additional examples. We have concentrated on the chat server example for exposition, but
have also implemented an extended chat server and a multiplayer game. These can be found

7 RELATED WORK

7.1 Session Types with Failure Handling

Carbone et al. [2008] provide the first formal basis for exceptions in a session-typed process calculus.
Our approach provides significant simplifications: zapper threads provide a simpler semantics and
remove the need for their queue levels, meta-reduction relation, and liveness protocol.
Our work draws on that of Mostrous and Vasconcelos [2014], who introduce the idea of cancellation. Our work differs from theirs in several key ways. Their system is a process calculus; ours is a $\lambda$-calculus. Their channels are synchronous; ours are asynchronous. Their exception handling construct scopes over a single action; ours scopes over an arbitrary computation.

Caires and Pérez [2017] describe a core, logically-inspired process calculus supporting non-determinism and abortable behaviours encoded via a nondeterminism modality. Processes may either provide or not provide a prescribed behaviour; if a process attempts to consume a behaviour that is not provided, then its linear continuation is safely discarded by propagating the failure of sessions contained within the continuation. Their approach is similar in spirit to our zapper threads. Additionally, they give a core $\lambda$-calculus with abortable behaviours and exception handling, and define a type-preserving translation into their core process calculus.

Our approach differs in several important ways. First, our semantics is asynchronous, handling the intricacies involved with cancelling values contained in message queues. Second, we give a direct semantics to EGV, whereas Caires and Pérez rely on a translation into their underlying process calculus. Third, to handle the possibility of disconnection, our calculus allows any channel to be discarded, whereas they opt for an approach more closely resembling checked exceptions, aided by a monadic presentation.

The above works are all theoretical. Backed by our theoretical development, our implementation integrates session types and exceptions, extending Links.

**Multiparty Session Types.** Fowler [2016] describes an Erlang implementation of the Multiparty Session Actor framework proposed by Neykova and Yoshida [2014, 2017b] with a limited form of failure recovery; Neykova and Yoshida [2017a] present a more comprehensive approach, based on refining existing Erlang supervision strategies. Chen et al. [2016] introduce a formalism based on multiparty session types [Honda et al. 2016] that handles partial failures by transforming programs to detect possible failures at a set of statically determined synchronisation points. These approaches rely on a fixed communication topology, using mechanisms such as dependency graphs or synchronisation points to determine which participants are affected when one participant fails. Delegation implies location transparency, thus we must consider dynamic topologies.
7.2 Session Types and Distribution

Hu et al. [2008] introduce Session Java (SJ), which allows distributed session-based communication in the Java programming language, making use of the Polyglot framework [Nystrom et al. 2003] to statically check session types. Hu et al. are the first to present the challenges of distributed delegation along with distributed algorithms which address those challenges. We adapt their algorithms to web applications. SJ restricts communication to a fixed set of simple types; Links allows arbitrary values to be sent. SJ provides statically scoped exception handling, propagating exceptions to ensure liveness (but this feature is not formalised).

Scalas and Yoshida [2016] introduce \texttt{\textbackslash l\textbackslash channel\textbackslash s}, a lightweight implementation of session types in Scala. To maximise applicability of their approach and not require any modifications to Scala, their approach detects duplicate endpoint use at runtime. By virtue of the translation into the linear \textit{\pi}-calculus introduced by Kobayashi [2003] and later expanded on by Dardha et al. [2017], \texttt{\textbackslash l\textbackslash channel\textbackslash s} is particularly amenable to distribution. Scalas et al. [2017] build upon this approach to translate a multiparty session calculus into the linear \textit{\pi}-calculus, providing the first distributed implementation of multiparty session types to support delegation.

7.3 Session Types via Affine Types

Rust [Matsakis and Klock II 2014] provides ownership types [Clarke 2003], ensuring that an object has at most one owner. Jespersen et al. [2015] use Rust’s ownership types to encode affine session types, but since affine endpoints can be discarded implicitly, their library does not guarantee progress. Although it is not possible to distinguish between dynamic failure and a developer forgetting to finish an implementation, our semantics can be implemented using Rust’s destructor mechanism, enabling a progress property [Kokke 2018].

8 CONCLUSION AND FUTURE WORK

Session types allow protocol conformance to be checked statically. The prevailing consensus has hitherto been to require that endpoints be used linearly to enforce session fidelity and prevent premature discarding of open channels. We have argued that in order to write realistic applications in the presence of distribution and failure, linearity should be supplemented with an \textit{explicit} cancellation operation. We show that, even in the presence of channel cancellation, our core calculus is well-behaved, being deadlock-free, type sound, confluent, and terminating.

In tandem with the formal development, we have developed an extension of the Links programming language to support distributed session-based communication for web applications, thus providing the first implementation of asynchronous session types with failure handling in a functional programming language. Our implementation leverages recent work on effect handlers.

Future work. Our implementation combines linearity and effect handlers. Linear effect handlers are new, and a ripe area of study in their own right; we plan to formalise session-typed concurrency and failure handling directly in terms of linear effect handlers. Multiparty session types [Honda et al. 2016] are yet to be included as a first-class construct of a core functional language. A natural starting point is to identify a \textit{\lambda}-calculus into which we can translate the MCP calculus of Carbone et al. [2016] and then investigate how our approach adapts to the multiparty setting.

ACKNOWLEDGMENTS

Thanks to James McKinna and the anonymous reviewers for detailed comments and suggestions. This work was supported by EPSRC grants EP/L01503X/1 (EPSRC CDT in Pervasive Parallelism) and EP/K034413/1 (From Data Types to Session Types—A Basis for Concurrency and Distribution), and an LFCS internship.
REFERENCES


Hillerström, Daniel and Sam Lindley. 2016. Liberating effects with rows and handlers. In TyDe@ICFP. ACM, 15–27.


## APPENDIX CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Example Runtime Typing Derivation</td>
<td>31</td>
</tr>
<tr>
<td>B</td>
<td>Deadlock-freedom</td>
<td>32</td>
</tr>
<tr>
<td>C</td>
<td>Supplement to Section 3 (Metatheory of EGV)</td>
<td>33</td>
</tr>
<tr>
<td>C.1</td>
<td>Preservation</td>
<td>33</td>
</tr>
<tr>
<td>C.2</td>
<td>Canonical Forms</td>
<td>46</td>
</tr>
<tr>
<td>C.3</td>
<td>Progress</td>
<td>46</td>
</tr>
<tr>
<td>C.4</td>
<td>Confluence</td>
<td>50</td>
</tr>
<tr>
<td>D</td>
<td>Supplement to Section 4.1 (Metatheory of EGV with Access Points)</td>
<td>51</td>
</tr>
<tr>
<td>E</td>
<td>Distributed Delegation</td>
<td>54</td>
</tr>
<tr>
<td>E.1</td>
<td>Challenges of Distributed Delegation</td>
<td>54</td>
</tr>
<tr>
<td>E.2</td>
<td>Approaches to Distributed Delegation</td>
<td>55</td>
</tr>
<tr>
<td>E.3</td>
<td>Delegation in Distributed Session Links</td>
<td>55</td>
</tr>
<tr>
<td>E.4</td>
<td>Correctness</td>
<td>56</td>
</tr>
</tbody>
</table>

A  EXAMPLE RUNTIME TYPING DERIVATION

We give an example derivation to illustrate how channels are introduced by name restrictions and then split into endpoints using the T-CONNECT₁ rules. We assume suitable encodings of linear booleans and integers using linear sums and products.

Let us assume we have derivations for:

\(\Gamma_1, a : !\text{Int.End} \vdash \circ E[\text{send} 5 a] : 1\)

\(\Gamma_2, b : ?\text{Bool}?\text{Int.End} \vdash E'[\text{receive} b] : A \quad \circ \vdash \text{true} : \text{Bool}\)

We construct a derivation \(D\) of \((v a)(v b)(\circ E[\text{send} 5 a] \parallel (a(e) \rightarrow \circ b(\text{true})) \parallel \bullet E'[\text{receive} b])\).

First let \(D_1\) be the following subderivation.

\[
\begin{align*}
\Gamma_2, b : ?\text{Bool}?\text{Int.End} & \vdash E'[\text{receive} b] : A \\
\Gamma_1, a : !\text{Int.End} & \vdash \circ E[\text{send} 5 a] \\
\text{T-BUFFER} &
\end{align*}
\]

Then let \(D_2\) be the following subderivation.

\[
\begin{align*}
\text{T-CONNECT₂} & \\
\Gamma_2, b : ?\text{Bool}?\text{Int.End} & \vdash E'[\text{receive} b] : A \\
\Gamma_1, a : !\text{Int.End} & \vdash \circ E[\text{send} 5 a] \\
D_1 &
\end{align*}
\]

The complete derivation \(D\) is as follows.

\[
\begin{align*}
\text{T-THREAD} & \\
\Gamma_1, a : !\text{Int.End} & \vdash \circ E[\text{send} 5 a] : 1 \\
\Gamma_1, a : !\text{Int.End} & \vdash \circ E[\text{send} 5 a] \\
D_2 &
\end{align*}
\]

\[
\begin{align*}
\text{T-CONNECT₁} & \\
\Gamma_1, \Gamma_2, a : (!\text{Int.End})^\circ, b : (?\text{Bool}?\text{Int.End})^\circ & \vdash \circ E[\text{send} 5 a] \parallel (a(e) \rightarrow \circ b(\text{true})) \parallel \bullet E'[\text{receive} b]) \\
\Gamma_1, \Gamma_2, a : (!\text{Int.End})^\circ & \vdash \circ E[\text{send} 5 a] \parallel (a(e) \rightarrow \circ b(\text{true})) \parallel \bullet E'[\text{receive} b]) \\
\text{T-NÜ} &
\end{align*}
\]

Let us read \(D\) bottom-upwards. The two instances of the T-NÜ rule introduce channels \(a\) and \(b\) into the runtime environment. The T-CONNECT₁ rule splits channel \(a\) into dual endpoints: on the left the endpoint \(a\) appears in the type environment and the sending thread; on the right the endpoint \(a\) appears in the runtime environment and the buffer. The T-CONNECT₂ rule splits channel \(b\) into dual endpoints: on the left the endpoint \(b\) appears in the runtime environment and the buffer; on the right the endpoint \(b\) appears in the type environment and the receiving thread.
B DEADLOCK-FREEDOM

Here we give a graph-theoretic account of deadlock-freedom in EGV, independent of our notion of progress, following Lindley and Morris [2015].

Due to the asynchronous semantics of EGV, sending on an endpoint and cancelling an endpoint reduce immediately. Deadlocks may therefore only occur when cycles occur receiving or closing an endpoint. We begin by classifying the notion of a blocked thread: that is, a thread which is blocked performing an action on some channel endpoint.

Definition B.1. We say that term \( M \) is blocked on name \( a \) if \( M \) is about to receive on or close \( a \). Formally:

\[
\text{blocked}(a, M) \triangleq \exists E. (M = E[\text{receive } a]) \lor (M = E[\text{close } a])
\]

Given the notion of a blocked thread, we may characterise the notion of a dependency between communication actions.

Definition B.2. Let \( C \) be a configuration such that \( a \) and \( b \) are not bound by \( C \). We say that \( a \) depends on \( b \) in \( C \), written \( \text{depends}(a, b, C) \), if \( C \) is a buffer connecting \( a \) and \( b \), or \( a \) appears in some thread blocked on \( b \), or if \( a \) depends on some name \( c \) which depends on \( b \). Formally:

- \( \text{depends}(a, b, a(\overrightarrow{V})) \triangleq \text{depends}(a, b, a(\overrightarrow{W})) \)
- \( \text{depends}(a, b, a(\overrightarrow{W})) \triangleq \text{depends}(a, b, a(\overrightarrow{V})) \)
- \( \text{depends}(a, b, \phi M) \triangleq \text{blocked}(b, M) \land a \in \text{fn}(M) \)
- \( \text{depends}(a, b, C) \triangleq \exists G, D, E, c. C \equiv G[D || E] \land \text{depends}(a, c, D) \land \text{depends}(c, b, E) \)

Remark. The above definition of dependency is an over-approximation to the intuitive notion, as a buffer need not have dependencies in both directions, but for our purposes this does not matter.

Definition B.3. We say that a configuration is deadlocked if it contains cyclic dependencies:

\[
\text{deadlocked}(C) \triangleq \exists D, E, a, b. C \equiv G[D || E] \land \text{depends}(a, b, D) \land \text{depends}(b, a, E)
\]

With these definitions in place, we can show that EGV configurations are deadlock-free.

Lemma B.4. If \( \text{depends}(a, b, C) \) then \( a, b \in \text{fn}(C) \).

Proof. By induction on the definition of \( \text{depends}(a, b, C) \).

\( \Box \)

Theorem B.5. If \( \Gamma ; \Delta \vdash C \), then \( \neg \text{deadlocked}(C) \).

Proof. By contradiction. Suppose \( \text{deadlocked}(C) \), that is:

\[
\exists D, E, a, b. C \equiv G[D || E] \land \text{depends}(a, b, D) \land \text{depends}(b, a, E)
\]

Thus by Lemma B.4, \( a, b \in \text{fn}(D) \) and \( b, a \in \text{fn}(E) \). Then by Lemma 3.6, \( C \) must be ill-typed. \( \Box \)

Remark. We regard blocked threads as deadlocked only if there is a cyclic dependency. It is perfectly possible for a configuration to include blocked threads without there being a deadlock.

- Deadlock-free open terms can block on external communication along a free endpoint.
- Deadlock-free closed terms can block on communication along an endpoint that appears in the return value of a program. This also amounts to being blocked on external communication.

All blocked threads can be ruled out by restricting the type of a program to be free of both session types and function types (the latter is necessary as closures can capture endpoints).
C SUPPLEMENT TO SECTION 3 (METATHEORY OF EGV)

C.1 Preservation

In this section, we present proofs that typeability is preserved by configuration reduction.

C.1.1 Equivalence. We begin by describing the properties of configuration equivalence. As described in §3, typeability of configurations is not preserved by equivalence. Nonetheless, Lemma C.1 shows that only the associativity of parallel composition may cause a configuration to be ill-typed.

Lemma C.1. If \( \Gamma; \Delta \vdash^\phi C \) and \( C \equiv D \), where the derivation of \( C \equiv D \) does not contain a use of the axiom for associativity, then \( \Gamma; \Delta \vdash^\phi D \).

Proof. By induction on the derivation of \( C \equiv D \), examining the equivalence in both directions to account for symmetry. We show that a typing derivation of the left-hand side of an equivalence rule implies the existence of the right-hand side, and vice versa.

That reflexivity, transitivity, and symmetry of the equivalence relation respect typing follows immediately because equality of typing derivations is an equivalence relation.

We make implicit use of the induction hypothesis.

Congruence rules

Case Name restriction

\[
\frac{C \equiv D}{(va)C \equiv (va)D}
\]

\[
\frac{\Gamma; \Delta, a : S^# \vdash^\phi C}{\Gamma; \Delta \vdash^\phi (va)C} \quad \text{iff} \quad \frac{\Gamma; \Delta, a : S^# \vdash^\phi D}{\Gamma; \Delta \vdash^\phi (va)D}
\]

Case Parallel Composition

\[
\frac{C \equiv D}{C \parallel E \equiv D \parallel E}
\]

There are three subcases, based on whether the parallel composition arises from T-CONNECT1, T-CONNECT2, or T-Mix.

Subcase T-Mix

\[
\frac{\Gamma_1; \Delta_1 \vdash^{\phi_1} C \quad \Gamma_2; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2 \vdash^{\phi_1 + \phi_2} C \parallel E} \quad \text{iff} \quad \frac{\Gamma_1; \Delta_1 \vdash^{\phi_1} D \quad \Gamma_2; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2 \vdash^{\phi_1 + \phi_2} D \parallel E}
\]

Subcase T-CONNECT1

\[
\frac{\Gamma_1, a : S; \Delta_1 \vdash^{\phi_1} C \quad \Gamma_2, a : S; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^# \vdash^{\phi_1 + \phi_2} C \parallel E} \quad \text{iff} \quad \frac{\Gamma_1, a : S; \Delta_1 \vdash^{\phi_1} D \quad \Gamma_2, a : S; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^# \vdash^{\phi_1 + \phi_2} D \parallel E}
\]

Subcase T-CONNECT2

\[
\frac{\Gamma_1; \Delta_1, a : S \parallel^{\phi_1} C \quad \Gamma_2, a : S; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^# \parallel^{\phi_1 + \phi_2} C \parallel E} \quad \text{iff} \quad \frac{\Gamma_1; \Delta_1, a : S \parallel^{\phi_1} D \quad \Gamma_2, a : S; \Delta_2 \vdash^{\phi_2} E}{\Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^# \parallel^{\phi_1 + \phi_2} D \parallel E}
\]
Equivalence Axioms

Case \( C \parallel D \equiv D \parallel C \)

There are three subcases, based on which rule is used for parallel composition.

Subcase T-Mix

\[
\Gamma_1; \Delta_1, \Delta_2, \phi_1, \phi_2, C \parallel D \quad \equiv \quad \Gamma_2; \Delta_2, \phi_2, \phi_1, \phi_2, D \parallel C
\]

Subcase T-Connect

\[
\Gamma_1, \Delta_1, \Delta_2, a : S, a : S \parallel \phi_1 \phi_2, C \parallel D \quad \equiv \quad \Gamma_2, \Delta_2, a : S, a : S \parallel \phi_1 \phi_2, D \parallel C
\]

Subcase T-Mix

\[
\Gamma_1; \Delta_1, \Delta_2, a : S, a : S \parallel \phi_1 \phi_2, C \parallel D \quad \equiv \quad \Gamma_2, \Delta_2, a : S, a : S \parallel \phi_1 \phi_2, D \parallel C
\]

Case \( C \parallel (va)D \equiv (va)(C \parallel D) \) if \( a \notin \text{fn}(C) \)

There are again three subcases based on which parallel composition rule is used. The exact rule does not affect the discussion, so without loss of generality we assume this is T-Mix.

\[
\Gamma_1; \Delta_1, \Delta_2, a : S^\# \parallel \phi_1 \phi_2, (va)D \quad \equiv \quad \Gamma_2; \Delta_2, a : S^\# \parallel \phi_1 \phi_2, (va)D
\]

In the left-to-right direction, that \( \Gamma_1, \Gamma_2; \Delta_1, \Delta_2, a : S^\# \) is well-defined follows because \( a \notin \text{fn}(C) \).

Case \((va)(vb)C \equiv (vb)(va)C\)

\[
\Gamma; \Delta, a : S^\#, b : T^\# \parallel \phi, C \quad \equiv \quad \Gamma; \Delta, a : S^\#, b : T^\# \parallel \phi, C
\]

Case \( a(\overrightarrow{V}) \sim b(\overrightarrow{W}) \equiv b(\overrightarrow{W}) \sim a(\overrightarrow{V}) \)

\[
\frac{S^\#/A = T^\#/B \quad \Gamma_1; \overrightarrow{V} : A \quad \Gamma_2; \overrightarrow{W} : B \quad \Gamma_1; \overrightarrow{V} \sim \overrightarrow{A}}{\Gamma_1, \Gamma_2; a : S, b : T^\# \quad a(\overrightarrow{V}) \sim b(\overrightarrow{W}) \quad \equiv \quad \Gamma_1, \Gamma_2; a : S, b : T^\# \quad b(\overrightarrow{W}) \sim a(\overrightarrow{V})}
\]
The above holds because $S/A = T/B \iff T/B = S/A:\n\nS/A = T/B \iff \text{(duality)}\nS/A = T/B \iff \text{(duality is involutive)}\nS/A = T/B \iff \text{(equality is symmetric)}\nT/B = S/A\n\nCase $\circ() \parallel C \equiv C$\n\n\n\nCase $(va)(vb)(\parallel a \parallel \parallel b \parallel a(e) \leftrightarrow b(e)) \parallel C \equiv C$\n\nWhile it is true that re-associating parallel composition may cause a configuration to be ill-typed, Lemma C.2 shows that it is always possible to re-associate parallel composition either directly, or by first commuting two subconfigurations.

**LEMMA C.2 (ASSOCIATIVITY).**

- If $\Gamma; \Delta \vdash^\phi C \parallel (D \parallel E)$, then either $\Gamma; \Delta \vdash^\phi (C \parallel D) \parallel E$ or $\Gamma; \Delta \vdash^\phi (C \parallel E) \parallel D$.
- If $\Gamma; \Delta \vdash^\phi (C \parallel D) \parallel E$, then either $\Gamma; \Delta \vdash^\phi C \parallel (D \parallel E)$ or $\Gamma; \Delta \vdash^\phi D \parallel (C \parallel E)$.

**PROOF.** The cases where either parallel composition arises by $\text{T-Mix}$ are unproblematic and can be re-associated without jeopardising typeability. Therefore, we concentrate on the cases where both compositions arise via $\text{T-Connect}$.

**Case $C \parallel (D \parallel E)$**
By the assumption that $\Gamma; \Delta \vdash^\phi C \parallel (D \parallel E)$ we have that $\Gamma = \Gamma_1, \Gamma_2, \Gamma_3$, and $\Delta = \Delta_1, \Delta_2, \Delta_3, a : S^\#, b : T^\#$, and $\phi = \phi_1 + \phi_2 + \phi_3$. There are 8 cases, based on whether $a, b \in \text{fn}(C)$ or $a, b \in \text{fn}(D)$ (it cannot be the case that $a, b \in \text{fn}(E)$, as $E$ only occurs under a single parallel composition), and the exact dualisation (i.e., whether composition happens via T-Connect$_1$ or T-Connect$_2$).

Of these, we are only interested in the cases where the sharing of the names differs, as opposed to the dualisation. Thus, we consider the following two cases, where both compositions occur using T-Connect$_1$:

(1) $\Gamma_1, a : S; \Delta_1 \vdash^\phi C$, and $\Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash^\phi_2 D$, and $\Gamma_3; \Delta_3, b : \overline{T} \vdash^\phi_3 E$

(2) $\Gamma_1, a : S; \Delta_1 \vdash^\phi C$, and $\Gamma_2, b : T; \Delta_2 \vdash^\phi_2 D$, and $\Gamma_3; \Delta_3, a : \overline{S}, b : \overline{T} \vdash^\phi_3 E$

**Subcase** $a \in \text{fn}(C), a, b \in D, b \in E$

\[
\begin{align*}
\Gamma_1, a : S; \Delta_1 \vdash^\phi C & \quad \Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash^\phi_2 D & \quad \Gamma_3; \Delta_3, b : \overline{T} \vdash^\phi_3 E \\
\Gamma_1, a : S; \Delta_1 \vdash^\phi C & \quad \Gamma_2, \Gamma_3; \Delta_2, a : \overline{S}, b : T^\# \vdash^\phi_1 + \phi_2 + \phi_3 D \parallel E & \quad \Gamma_3; \Delta_3, b : \overline{T} \vdash^\phi_3 E \\
\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, a : S^\#, b : T^\# \vdash^\phi_1 + \phi_2 + \phi_3 (C \parallel D) \parallel E
\end{align*}
\]

As $D$ contains both $a$ and $b$, associativity does not alter the sharing of names and may be applied safely.

\[
\begin{align*}
\Gamma_1, a : S; \Delta_1 \vdash^\phi C & \quad \Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash^\phi_2 D \\
\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, a : S^\#, b : T^\# \vdash^\phi_1 + \phi_2 + \phi_3 (C \parallel D) \parallel E
\end{align*}
\]

**Subcase** $a \in \text{fn}(C); b \in D; a, b \in E$

\[
\begin{align*}
\Gamma_1, a : S; \Delta_1 \vdash^\phi C & \quad \Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash^\phi_2 D & \quad \Gamma_3; \Delta_3, a : \overline{S}, b : \overline{T} \vdash^\phi_3 E \\
\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, a : S^\#, b : T^\# \vdash^\phi_1 + \phi_2 + \phi_3 C \parallel (D \parallel E)
\end{align*}
\]

Here, we may not apply associativity directly. But, we may first commute $D$ and $E$:

\[
\begin{align*}
\Gamma_1, a : S; \Delta_1 \vdash^\phi C & \quad \Gamma_2, \Gamma_3; \Delta_2, a : \overline{S}, b : \overline{T} \vdash^\phi_3 E \\
\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, a : S^\#, b : T^\# \vdash^\phi_1 + \phi_2 + \phi_3 (C \parallel D) \parallel E
\end{align*}
\]

and from here we may safely re-associate to the left:

\[
\begin{align*}
\Gamma_2, a : S; \Delta_2 \vdash^\phi C & \quad \Gamma_3; \Delta_3, a : \overline{S}, b : \overline{T} \vdash^\phi_3 E \\
\Gamma_2, \Gamma_3; \Delta_2, a : \overline{S}, b : \overline{T} \vdash^\phi_1 + \phi_2 + \phi_3 (C \parallel D) \parallel E
\end{align*}
\]

**Case** $(C \parallel D) \parallel E$

(1) $\Gamma_1, a : S; \Delta_1 \vdash^\phi C$, and $\Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash^\phi_2 D$, and $\Gamma_3; \Delta_3, b : \overline{T} \vdash^\phi_3 E$

(2) $\Gamma_1, a : S, b : T; \Delta_1 \vdash^\phi C$, and $\Gamma_2, \Delta_2, b : \overline{T} \vdash^\phi_2 D$, and $\Gamma_3, \Delta_3, a : \overline{S} \vdash^\phi_3 E$

**Subcase** $a \in C; a, b \in D; b \in E$
Application associativity here does not make the configuration ill-typed, as $\mathcal{D}$ contains both names:

\[
\frac{\Gamma_1, a : S; \Delta_1 \vdash \phi_1 C \quad \Gamma_2, b : T; \Delta_2, a : \overline{S} \vdash \phi_2 C \quad \mathcal{D} \quad \Gamma_3; \Delta_3, b : \overline{T} \vdash \phi_3 E}{\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, \Delta_3, a : S\# b : T\# \overline{\phi_1 + \phi_2 + \phi_3} \mathcal{C} \parallel \mathcal{D} \parallel E}
\]

Subcase $a, b \in C; a \in \mathcal{D}; b \in \mathcal{E}$

Assumption:

\[
\frac{\Gamma_1, a : S, b : T; \Delta_1 \vdash \phi_1 C \quad \Gamma_2, \Delta_2, a : \overline{S} \vdash \phi_2 C \quad \mathcal{D} \quad \Gamma_3; \Delta_3, b : \overline{T} \vdash \phi_3 E}{\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, \Delta_3, a : S\# b : T\# \overline{\phi_1 + \phi_2 + \phi_3} \mathcal{C} \parallel \mathcal{D} \parallel \mathcal{E}}
\]

By commutativity:

\[
\frac{\Gamma_2; \Delta_3, b : \overline{T} \vdash \phi_3 E \quad \Gamma_1, a : S, b : T; \Delta_1 \vdash \phi_1 C}{\Gamma_2; \Delta_3, b : \overline{T} \vdash \phi_3 E}
\]

By associativity:

\[
\frac{\Gamma_1, a : S, b : T; \Delta_1 \vdash \phi_1 C \quad \Gamma_3; \Delta_3, b : \overline{T} \vdash \phi_3 E}{\Gamma_1, \Gamma_2, \Gamma_3; \Delta_1, \Delta_2, \Delta_3, a : S\# b : T\# \overline{\phi_1 + \phi_2 + \phi_3} \mathcal{C} \parallel \mathcal{D} \parallel \mathcal{E}}
\]

as required.

\[\square\]

C.1.2 Configuration Reduction. We may now show that configuration reduction preserves typeability of configurations. We begin by stating some auxiliary results about substitution and evaluation contexts.

Typing of terms is preserved by substitution.

Lemma C.3 (Substitution). If:

\begin{enumerate}
\item $\Gamma \vdash M : B$
\item $\Gamma, x : B \vdash N : A$
\item $\Gamma_1, \Gamma_2$ is well-defined
\end{enumerate}

then $\Gamma_1, \Gamma_2 \vdash N\{M/x\} : A$.

Proof. By induction on the derivation of $\Gamma_2, x : B \vdash N : A$. \[\square\]

Lemma C.4 shows that a subterm of a well-typed evaluation context $E$ (and therefore also a pure evaluation context $P$) is typeable with a subset of the type environment. Lemma C.5 states that the subterm of a well-typed evaluation context can be replaced. Both follow the formulation of Gay and Vasconcelos [2010].
Lemma C.4 (Typability of Subterms). If $D$ is a derivation of $\Gamma \vdash E[M] : A$, then there exist $\Gamma_1, \Gamma_2$ and $B$ such that $\Gamma = \Gamma_1, \Gamma_2$, that $D$ has a subderivation $D'$ that concludes $\Gamma_2 \vdash M : B$, and the position of $D'$ in $D$ corresponds to the position of the hole in $E$.


Lemma C.5 (Replacement (evaluation contexts)). If:

- $D$ is a derivation of $\Gamma_1, \Gamma_2 \vdash E[M] : A$
- $D'$ is a subderivation of $D$ concluding $\Gamma_2 \vdash M : B$
- The position of $D'$ in $D$ corresponds to that of the hole in $E$
- $\Gamma_3 \vdash N : B$
- $\Gamma_1, \Gamma_3$ is well-defined

then $\Gamma_1, \Gamma_3 \vdash E[N] : A$.


To prove preservation on configurations, we must first establish some auxiliary results on configuration contexts. Lemma C.6 states how we may type subconfigurations.

Lemma C.6 (Typability of Subconfigurations). If $D$ is a derivation of $\Gamma; \Delta \vdash^\phi \mathcal{G}[C]$, then there exist $\Gamma'; \Delta', \phi'$ such that $D$ has a subderivation $D'$ that concludes $\Gamma'; \Delta' \vdash^\phi' C$, and the position of $D'$ in $D$ corresponds to the position of the hole in $\mathcal{G}$.

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

Lemma C.7 states that we may replace a subconfiguration of a configuration context. The lemma is slightly complicated by the fact that $(va)\mathcal{G}$ binds a variable $a$, but replacement is safe if the typing environments are related by the environment reduction relation.

Lemma C.7 (Replacement (configurations)). If:

- $D$ is a derivation of $\Gamma; \Delta \vdash^\phi \mathcal{G}[C]$
- $D'$ is a subderivation of $D$ concluding that $\Gamma'; \Delta' \vdash^\phi' C$ for some $\Gamma', \Delta', \phi'$
- $\Gamma''; \Delta'' \vdash^\phi' C'$ for some $\Gamma'', \Delta''$ such that $\Gamma''; \Delta'' \rightarrow^2 \Gamma'''; \Delta'''$
- The position of $D$ in $D'$ corresponds to that of the hole in $\mathcal{G}$

then there exist some $\Gamma''', \Delta'''$ such that $\Gamma'''; \Delta''' \vdash^\phi \mathcal{G}[C']$ and $\Gamma; \Delta \rightarrow^2 \Gamma'''; \Delta'''$.

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

Theorem 3.2 (Preservation (Configurations))

Assume $\Gamma$ only contains entries of the form $a_i : S_i$.

If $\Gamma; \Delta \vdash^\phi C$ and $C \rightarrow D$, then there exist $\Gamma', \Delta'$ such that $\Gamma; \Delta \rightarrow^2 \Gamma'; \Delta'$ and $\Gamma'; \Delta' \vdash^\phi D$.

Proof. By induction on the derivation of $C \rightarrow D$. Where there is a choice of value for $\phi$, we consider the case where $\phi = \bullet$; the cases where $\phi = \circ$ are similar.

Case E-Fork

Assumption:

$$\Gamma_1, \Gamma_2 \vdash \bullet E[\text{fork } \lambda x.M] : A$$

$$\Gamma_1, \Gamma_2, \vdash^* \bullet E[\text{fork } \lambda x.M]$$
By Lemma C.4:

\[
\Gamma_2, x : S \vdash M : 1 \\
\Gamma_2 \vdash \lambda x. M : S \rightarrow \circ 1 \\
\Gamma_2 \vdash \text{fork} \lambda x. M : S
\]

By Lemma C.3, \(\Gamma_2, b : S \vdash M(b/x) : 1\), and by Lemma C.5, \(\Gamma_1, a : \overline{S} \vdash E[a] : A\). As duality is involutive, \(\overline{S} = S\).

Reconstructing:

\[
S/\epsilon = \overline{S}/\epsilon \quad \cdot \epsilon : \epsilon \quad \cdot \epsilon : \epsilon
\]

\[
\Gamma_1, a : \overline{S} \vdash E[a] : A \\
\Gamma_2, b : S \vdash \circ M(b/x) \\
\Gamma_1, a : \overline{S}; \vdash^* E[a] \\
\Gamma_2, a : S, b : S^\# \vdash \circ M(b/x) ; (a(e) \sim b(e))
\]

**Case E-Send**

Assumption:

\[
\Gamma_1, \Gamma_2, a : S^\# \cdot \vdash^* E[a] ; (vb)(E[a]) \quad \circ M(b/x) ; (a(e) \sim b(e))
\]

\[
\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, a : S^\# ; b : T \cdot \vdash^* E[a] ; (vb)(E[a]) \quad (a(e) \sim b(e))
\]

By Lemma C.4:

\[
\Gamma_2 \vdash U : A \quad a : !A.S' \vdash a : !A.S'
\]

Thus, \(S = !A.S'\), and \(\overline{S} = ?A.S'.\) We may therefore refine our original derivation:

\[
\Gamma_1, \Gamma_2, a : !A.S' + E[\text{send} U a] : C \\
\Gamma_3, \Gamma_4, a : \overline{S}; b : T \cdot \vdash^* E[\text{send} U a] ; (a(e) \sim b(e))
\]

Since \(?A.S'/A = T/B\) is well-defined, we have that \(\overline{A} = \epsilon\). By the definition of slicing, we have that \(\overline{B} = B_1, \ldots, B_n\). It follows that \(\overline{S'/A} = T/B \cdot A\).

By Lemma C.5, we have \(\Gamma_1, \Gamma_2, a : S' \vdash E[a] : C\).

Reconstructing:

\[
\Gamma_1, a : S' \vdash E[a] : C \\
\Gamma_2, b : S^\# \vdash \circ M(b/x) \\
\Gamma_1, a : S'; \vdash^* E[a] \\
\Gamma_2, \Gamma_3, \Gamma_4, a : \overline{S}; b : T \cdot \vdash^* E[\text{send} U a] ; (a(e) \sim b(e))
\]

Finally, we must show environment reduction:

\[ !A.S' \rightarrow S' \]

\[ \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : !A.S') \parallel b : T \rightarrow \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : S', b : T \]

as required.

Case E-Receive

Assumption:

\[ \Gamma_1, a : S \vdash E[\text{receive } a] : C \quad \overline{S/A} = \overline{T/B} \quad \Gamma_2, \Gamma_3 + U \cdot \overline{V} : \overline{A} \quad \Gamma_4 + \overline{W} : B \]

\[ \Gamma_1, a : S; \vdash^* E[\text{receive } a] \quad \Gamma_2, \Gamma_3, \Gamma_4; a : \overline{S}, b : T \vdash^* a(U \cdot \overline{V}) \leadsto b(\overline{W}) \]

\[ \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : !A.S') \parallel b : T \vdash^* E[\text{receive } a] \parallel a(U \cdot \overline{V}) \leadsto b(\overline{W}) \]

By Theorem C.4:

\[ a : ?A.S' \vdash a : ?A.S' \]

\[ a : ?A.S' \vdash \text{receive } a : (A \times S') \]

Thus, we have that \( S = ?A.S' \) and \( \overline{S} = !A.S' \), and we may therefore refine the original typing derivation:

\[ \Gamma_1, a : S' \vdash E[(U, a)] : C \quad \overline{S'}/A' = \overline{T'/B} \quad \Gamma_2, \Gamma_3 + U \cdot \overline{V} : A \cdot \overline{A'} \quad \Gamma_4 + \overline{W} : B \]

\[ \Gamma_1, a : S'; \vdash^* E[(U, a)] \quad \Gamma_2, \Gamma_3, \Gamma_4; a : !A.S', b : T \vdash^* a(U \cdot \overline{V}) \leadsto b(\overline{W}) \]

\[ \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : (\overline{S'}) \parallel b : T \vdash^* E[(U, a)] \parallel a(\overline{V}) \leadsto b(\overline{W}) \]

By Lemma C.5, we have \( \Gamma_1, \Gamma_2, a : S' \vdash E[(U, a)] : C \) (that \( \Gamma_1, \Gamma_2 \) is defined follows from the fact that \( \Gamma_1 \) and \( \Gamma_2 \) are sub-environments of the original typing environment and are therefore necessarily disjoint).

By the definition of slicing, \( !A.S'/A \vdash \overline{A'} \iff \overline{S'}/A' \).

Thus, recomposing:

\[ \Gamma_1, \Gamma_2, a : S' \vdash E[(U, a)] : C \quad \overline{S'/A'} = \overline{T'/B} \quad \Gamma_2, \Gamma_3 + \overline{V} : \overline{A'} \quad \Gamma_4 + \overline{W} : B \]

\[ \Gamma_1, \Gamma_2, a : S'; \vdash^* E[(U, a)] \quad \Gamma_3, \Gamma_4; a : \overline{S'}, b : T \vdash^* a(\overline{V}) \leadsto b(\overline{W}) \]

\[ \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : S', b : T \vdash^* E[(U, a)] \parallel a(\overline{V}) \leadsto b(\overline{W}) \]

Finally, we must show environment reduction:

\[ ?A.S' \rightarrow S' \]

\[ \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : (?A.S') \parallel b : T \rightarrow \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4; a : S', b : T \]

as required.

Case E-Close
where $fn(U) = \{c_i\}_i$. 

Assumption:

\[
S/A = T/B \quad \Gamma_1, \Gamma_2 \vdash U \cdot V : A \quad \Gamma_3 \vdash W : B
\]

\[
a : S; \vdash o \vdash a
\]

\[
\Gamma_1, \Gamma_2, \Gamma_3; a : S', b : T \vdash o \vdash a(U \cdot V) \sim \sim b(W)
\]

By the definition of slicing, we have that there exist some \(A\) and \(S'\) such that \(S = !A.S'\). Thus, we may refine our judgement:

\[
a : ?A.S'; \vdash o \vdash a
\]

\[
\Gamma_1, \Gamma_2, \Gamma_3; a : S', b : T \vdash o \vdash a(U \cdot V) \sim \sim b(W)
\]

By the definition of buffer typing, we have that \(\Gamma_1 \vdash U : A\). By the definition of the reduction rule, \(f_n(U) = \{c_i\}_i\), and by assumption, \(\Gamma\) contains only runtime names. Thus, we may conclude that \(U\) is closed and therefore that \(\Gamma_1 = c_1 : S_1, \ldots, c_n : S_n\) for some session types \(S_1, \ldots, S_n\).

By the definition of slicing, we have that \(!A.S'/A' \equiv \equiv S'/A'\). Correspondingly, by T-Buffer, we may show

\[
S'/A' = T/B \quad \Gamma_2 \vdash V : A' \quad \Gamma_3 \vdash W : B
\]

By repeated applications of T-Zap and T-Mix, we have that

\[
\Gamma_2, \Gamma_3, c_1 : S_1, \ldots, c_n : S_n; a : S', b : T \vdash o \vdash c_1 || \ldots || c_n || a(V) \sim \sim b(W)
\]

Recomposing:

\[
c_n : S_n; \vdash o \vdash c_n
\]

\[
\Gamma_2, \Gamma_1; a : S', b : T \vdash o \vdash a(V) \sim \sim b(W)
\]

\[
a : S'; \vdash o \vdash a
\]

\[
\Gamma_2, \Gamma_3, c_1 : S_1, \ldots, c_n : S_n; a : S', b : T \vdash o \vdash c_1 || \ldots || c_n || a(V) \sim \sim b(W)
\]

Finally, we must show environment reduction:

\[
?A.S' \longrightarrow S'
\]

\[
\Gamma_2, \Gamma_3, c_1 : S_1, \ldots, c_n : S_n; a : (?A.S'^d), b : T \longrightarrow \Gamma_2, \Gamma_3, c_1 : S_1, \ldots, c_n : S_n; a : S'^d, b : T
\]

as required.

Case E-CloseZap

\[
\mathcal{F}[\text{close } a] \parallel \vdash b \parallel a(e) \sim \sim b(e) \longrightarrow \mathcal{F}[\text{raise}] \parallel \vdash a \parallel \vdash b \parallel a(e) \sim \sim b(e)
\]
Assumption:

\[ \Gamma, a : S \vdash E[\text{close} \ a] : C \]

\[ b : T; \vdash \circ \ b \quad \vdash : \overline{S}, b : \overline{T} \vdash \epsilon \ a(\epsilon) \leadsto b(\epsilon) \]

\[ \Gamma ; a : S\# , b : T\# \vdash \ast E[\text{close} \ a] \parallel \circ b \parallel a(\epsilon) \leadsto b(\epsilon) \]

By Lemma C.4:

\[ a : \text{End} \vdash a : \text{End} \]

\[ a : S \vdash \text{close} \ a : 1 \]

We may therefore refine our original derivation:

\[ \Gamma , a : \text{End} \vdash E[\text{close} \ a] : C \]

\[ \text{End} = \text{End} \quad \vdash \epsilon : \epsilon \quad \vdash : \epsilon \]

\[ b : \text{End} ; \vdash \circ b \quad \vdash : \overline{a}, b : \overline{\text{End}} \vdash \ast a(\epsilon) \leadsto b(\epsilon) \]

\[ \Gamma ; a : \text{End}\# , b : \text{End}\# \vdash \ast E[\text{close} \ a] \parallel \circ b \parallel a(\epsilon) \leadsto b(\epsilon) \]

By Lemma C.5, \( \Gamma \vdash E[\text{raise}] : C \).

Thus, recomposing:

\[ \Gamma \vdash E[\text{raise}] : C \]

\[ a : \text{End} ; \vdash \circ a \quad \vdash : \overline{a}, b : \overline{\text{End}} \vdash \ast a(\epsilon) \leadsto b(\epsilon) \]

\[ \Gamma ; a : \text{End}\# , b : \text{End}\# \vdash \ast E[\text{close} \ a] \parallel \circ b \parallel a(\epsilon) \leadsto b(\epsilon) \]

as required.

**Case E-ReceiveZap**

\[ \ast E[\text{receive} \ a] \parallel \circ b \parallel a(\epsilon) \leadsto b(\overline{W}) \rightarrow \ast E[\text{raise}] \parallel \circ a \parallel \circ b \parallel a(\epsilon) \leadsto b(\overline{W}) \]

Assumption:

\[ \overline{S}/\epsilon = \overline{T}/\overline{B} \quad \vdash : \epsilon \quad \Gamma_2 \vdash \overline{W} : \overline{B} \]

\[ \Gamma_1 , a : S \vdash E[\text{receive} \ a] : C \]

\[ b : T; \vdash \circ b \quad \Gamma_2 ; a : \overline{S}, b : \overline{T} \vdash \ast a(\epsilon) \leadsto b(\overline{W}) \]

\[ \Gamma_1 ; a : \text{End}, b : \text{End} \vdash \ast E[\text{receive} \ a] \parallel \circ b \parallel a(\epsilon) \leadsto b(\overline{W}) \]

By Lemma C.4:

\[ a : \overline{S}/\epsilon \vdash a : \overline{S}/\epsilon \]

\[ a : \overline{S}/\epsilon \vdash \text{receive} \ a : (A \times S') \]
By Lemma C.5, \( \Gamma_1 \vdash E[\text{raise}] : S' \). Thus, recomposing:

\[
\begin{array}{c}
\Gamma_1 \vdash E[\text{raise}] : C \\
\Gamma_1 ; \cdot \vdash^* E[\text{raise}]
\end{array}
\]

\[
\begin{array}{c}
\Gamma_2 ; a : S, b : T \vdash^0 a(e) \leadsto b(W)
\end{array}
\]

as required.

**Case E-Raise**

\[\bullet E[\text{try } P[\text{raise}] \text{ as } x \text{ in } M \text{ otherwise } N] \rightarrow E[N] \parallel \frac{\cdot c_1 \parallel \cdots \parallel \cdot c_n}{\text{and } \text{fn}(P) = \{c_i\}_i.}

\text{Assumption:}

\[\Gamma \vdash E[\text{try } P[\text{raise}] \text{ as } x \text{ in } M \text{ otherwise } N] : A' \]

\[\Gamma_1 ; \cdot \vdash^* E[\text{try } P[\text{raise}] \text{ as } x \text{ in } M \text{ otherwise } N]
\]

By Lemma C.4, there exist \( \Gamma_1, \Gamma_2, A, B, C \) such that \( \Gamma = \Gamma_1, \Gamma_2, \Gamma_3 \) and

\[\Gamma_2 \vdash P[\text{raise}] : A \quad \Gamma_3, x : B \vdash M : C \quad \Gamma_3 \vdash N : C \]

\[\Gamma_2, \Gamma_3 \vdash \text{try } P[\text{raise}] \text{ as } x \text{ in } M \text{ otherwise } N : C
\]

Since \( \Gamma \) contains only runtime names and \( \text{fn}(P) = \{c_i\}_i \), we know that \( \Gamma_2 = c_1 : S_1, \ldots, c_n : S_n \) for some \( S_i \).

By Lemma C.5, we have that:

\[\Gamma_1, \Gamma_3 \vdash E[N] : A'
\]

By repeated applications of T-ZAP and T-MIX, we have that \( \Gamma_2 \parallel \frac{\cdot c_1 \parallel \cdots \parallel \cdot c_n}{\text{Therefore, recomposing:}}

\[\Gamma_1, \Gamma_3 ; \cdot \vdash^* E[N] \]

\[
\begin{array}{c}
\frac{c_1 : S_1; \cdot \vdash^0 \cdot c_1}{c_1 : S_1, \ldots, c_n : S_n; \cdot \vdash^0 \frac{\cdot c_1 \parallel \cdots \parallel \cdot c_n}{\text{as required.}}
\end{array}
\]

**Case E-RaiseChild**

\[\circ P[\text{raise}] \rightarrow \frac{\cdot c_1 \parallel \cdots \parallel \cdot c_n}{\text{Assumption:}}

\[\Gamma \vdash P[\text{raise}] : 1 \]

\[\frac{\cdot \vdash^0 \circ P[\text{raise}]}{}
\]
By Lemma C.4, the knowledge that $\Gamma$ contains only runtime names, the knowledge that $\text{fn}(P) = \langle c_1, \ldots, c_n \rangle$, and the typing rule $T\text{-RAISE}$, we have that $\Gamma = c_1 : S_1, \ldots, c_n : S_n$ for some session types \{\text{S}_i\}_i.

Thus, by repeated applications of $T\text{-ZAP}$ and $T\text{-MIX}$, we may deduce that

$$\Gamma ; \vdash \begin{array}{c} c_1 \parallel \cdots \parallel c_n \end{array}$$

as required.

**Case** E-RaiseMain

$$\bullet P[\text{raise}] \rightarrow \text{halt} \parallel \begin{array}{c} \vdash c_1 \parallel \cdots \parallel c_n \end{array}$$

where $\text{fn}(P) = \langle c_i \rangle_i$.

**Assumption:**

$$\Gamma \vdash P[\text{raise}] : C$$

$$\Gamma ; \vdash \bullet P[\text{raise}]$$

By Lemma C.4, the knowledge that $\Gamma$ contains only runtime names, the knowledge that $\text{fn}(P) = c_1, \ldots, c_n$, and the typing rule $T\text{-RAISE}$, we have that $\Gamma = c_1 : S_1, \ldots, c_n : S_n$ for some session types \{\text{S}_i\}_i.

By repeated applications of $T\text{-ZAP}$ and $T\text{-MIX}$, we may deduce that

$$\Gamma ; \vdash \begin{array}{c} \vdash c_1 \parallel \cdots \parallel c_n \end{array}$$

By $T\text{-HALT}$, we have that $\vdash \begin{array}{c} \vdash \text{halt} \end{array}$. Thus, recomposing, we arrive at

$$\Gamma_1, \Gamma_3, c_1 : S_1, \ldots, c_n : S_n ; \vdash \begin{array}{c} \vdash c_1 \parallel \cdots \parallel c_n \end{array}$$

as required.

**Case** LiftC

**Assumptions:**

- $\Gamma ; \Delta \vdash^\phi G[C]$.
- $C \rightarrow D$

Let $D$ be a derivation of $\Gamma ; \Delta \vdash^\phi G[C]$. By Lemma C.6, we have that there exists some $D'$ such that $D'$ is a subderivation of $D$ concluding $\Gamma' ; \Delta' \vdash^\phi C$, where the position of $D'$ in $D$ corresponds to that of the hole in $G$.

By the IH, we have that there exists some $\Gamma'' ; \Delta''$ such that $\Gamma ; \Delta \rightarrow^? \Gamma'' ; \Delta''$ and $\Gamma'' ; \Delta'' \vdash^\phi D$. By Lemma C.7, we have that there exist some $\Gamma''' ; \Delta'''$ such that $\Gamma ; \Delta \rightarrow^? \Gamma''' ; \Delta'''$ and $\Gamma''' ; \Delta''' \vdash^\phi G[D]$, as required.

**Case** E-LiftM

**Assumptions:**

$$\Gamma \vdash M : A$$

$$\Gamma ; \vdash^\bullet \bullet M$$
and $M \rightarrow_M N$. By Lemma 3.1, we have that $\Gamma \vdash N : A$. Recomposing:

\[
\begin{align*}
\Gamma \vdash N : A \\
\Gamma ; \vdash \bullet {\cdot}N
\end{align*}
\]
as required. $\square$

### C.2 Canonical Forms

**Theorem 3.7: Canonical Forms** Given $C$ such that $\Gamma ; \Delta \vdash \bullet C$, there exists some $C' \equiv C$ such that $\Gamma ; \Delta \vdash \bullet C'$ and $C'$ is in canonical form.

**Proof.** The proof is by induction on the count of $\nu$-bound variables, following Lindley and Morris [2015]. Without loss of generality, assume that the $\nu$-bound variables of $C$ are distinct. Let \{ai | 1 ≤ i ≤ n\} be the set of $\nu$-bound variables in $C$ and let \{Dj | 1 ≤ j ≤ m\} be the set of threads in $C$.

In the case that $n = 0$, by Lemma C.1 we can safely commute the main thread such that it is the rightmost configuration, and associate parallel composition to the right using Lemma C.2 to derive a well-typed canonical form.

In the case that $n \geq 1$, pick some ai and Dj such that ai is the only $\nu$-bound variable in fn(Dj); Lemma 3.6 and a standard counting argument ensure that such a name and configuration exist. By the equivalence rules, there exists $E$ such that $\Gamma ; \Delta \vdash \phi C \equiv (\nu a_i)(D_j \parallel E)$ (that ai is the only $\nu$-bound variable in fn(Dj) ensures well-typing). Moreover, we have that there exist $\Gamma' \subseteq \Gamma$, $\Delta' \subseteq \Delta$, and $S$ such that either $\Gamma' \vdash S; \Delta' \vdash \phi E$ or $\Gamma'; \Delta', ai : S \vdash \phi E$. By the induction hypothesis, there exists $E'$ in canonical form such that either $\Gamma', ai : S; \Delta' \vdash \phi E \equiv E'$ or $\Gamma'; \Delta', ai : S \vdash \phi E \equiv E'$. Let $C' = (\nu a_i)(D_j \parallel E')$. By construction it holds that $\Gamma; \Delta \vdash \phi C \equiv C'$ and that $C'$ is in canonical form. $\square$

### C.3 Progress

To prove Theorem 3.9, we prove a similar property in which canonical configurations are decomposed step-by-step rather than in one go.

**Definition C.8 (Open Progress).** Suppose $\Psi; \Delta \vdash \bullet C$, where $C$ is in canonical form and $C \Rightarrow$. We say that $C$ satisfies open progress if:

1. $C = (\nu a)(A \parallel D)$, where $\Psi = \Psi_1, \Psi_2$ and $\Delta = \Delta_1, \Delta_2$ such that either:
   a. $\Psi_1, a : S; \Delta_1 \vdash A$ and $\Psi_2; \Delta_2, a : \overline{S} \vdash D$ where $D$ satisfies open progress, and $A$ is either:
      i. A thread \( \circ M \) where ready($b, M$) for some $b \in \text{fn}(\Psi_1, a : S)$; or
      ii. A zapper thread \( \z a \); or
      iii. A buffer $b(\overline{V}) \Rightarrow_c c(\overline{W})$ where $b, c \neq a$ and either $a \in \overline{V}$ or $a \in \overline{W}$
   b. $\Psi_1; \Delta_1, a : \overline{S} \vdash A$ and $\Psi_2, a : S; \Delta_2 \vdash \bullet D$, where $D$ satisfies open progress, and $A$ is either:
      a. $a(\overline{V}) \Rightarrow_b b(\overline{W})$ or $b(\overline{V}) \Rightarrow_a a(\overline{W})$ for some $b \in \text{fn}(\Delta_1)$
2. $C = A \parallel M$, where $\Psi = \Psi_1, \Psi_2$ and either:
   a. $\Delta = \Delta_1, \Delta_2, a : S^\#$, where $\Psi_1, a : S; \Delta_1 \vdash A$ and $\Psi_2; \Delta_2, a : \overline{S} \vdash \bullet M$, where $M$ satisfies open progress, and $A$ is either:
      i. A thread $\circ M$ where ready($b, M$) for some $b \in \text{fn}(\Psi_1, a : S)$; or
      ii. A zapper thread $\z a$; or
   b. $\Delta = \Delta_1, \Delta_2, a : S^\#$, where $\Psi_1; \Delta_1, a : \overline{S} \vdash A$ and $\Psi_2, a : S; \Delta_2 \vdash \bullet M$, where $M$ satisfies open progress, and $A$ is either:
      a. $a(\overline{V}) \Rightarrow_b b(\overline{W})$ or $b(\overline{V}) \Rightarrow_a a(\overline{W})$ for some $b \in \text{fn}(\Delta_1)$
(c) \( \Delta = \Delta_1, \Delta_2 \), where \( \Psi_1; \Delta_1 \vdash^{\phi} \mathcal{A} \) and \( \Psi_2; \Delta_2 \vdash^{\bullet} M \), where \( M \) satisfies open progress, and \( \mathcal{A} \) is either:

(i) A thread \( \circ M \) where either \( M = () \), or ready\((a, M)\) for some \( a \in \text{fn}(\Psi_1) \); or

(ii) A zapper thread \( \triangledown a \) for some \( a \in \text{fn}(\Psi_1) \); or

(iii) A buffer \( a(\overrightarrow{V}) \leftrightarrow b(\overrightarrow{W}) \) for some \( a, b \in \text{fn}(\Delta_1) \)

1. \( C = \mathcal{T} \), where either:

(a) \( \mathcal{T} = \bullet N \), where \( N \) is either a value or ready\((b, N)\) for some \( b \in \text{fn}(\Psi) \)

(b) \( \mathcal{T} = \mathit{halt} \)

**Lemma C.9.** Suppose \( \Psi; \Delta \vdash^{\bullet} C \), where \( C \) is in canonical form and \( C \Rightarrow \). Then \( C \) satisfies open progress.

**Proof.** By induction on the derivation of \( \Psi; \Delta \vdash^{\bullet} C \). We have three cases, based on the structure of the given canonical form.

**Case** \( C = (va)(\mathcal{A} \parallel D) \), with \( a \in \text{fn}(\mathcal{A}) \), and where \( D \) is in canonical form

By assumption, we know that \( \Psi; \Delta \vdash^{\phi} (va)(\mathcal{A} \parallel D) \).

This configuration is typeable by T-Nu, followed by either T-CONNECT\(_1\) or T-CONNECT\(_2\). As the definition of canonical forms requires that \( a \in \text{fn}(\mathcal{A}) \), it cannot be the case that the parallel composition arises as a result of T-Mix.

We consider these two subcases to show that \( \mathcal{A} \) satisfies the properties required by open progress.

**Subcase** T-CONNECT\(_1\)

\[
\Psi_1, a : S; \Delta_1 \vdash^{\phi_1} \mathcal{A} \quad \Psi_2, a : S; \Delta_2 \vdash^{\phi_2} D
\]

\[
\Psi_1, \Psi_2; \Delta_1, \Delta_2, a : S^\# \vdash^{\phi_1 + \phi_2} \mathcal{A} \parallel D
\]

By the definition of auxiliary threads and inversion on the typing relation, we know that \( \mathcal{A} \) is of the following forms:

- \( \circ M \), where \( a \in \text{fn}(M) \), and \( \Psi_1, a : S \vdash M : 1 \)
- \( \triangledown a \)
- \( b(\overrightarrow{V}) \leftrightarrow c(\overrightarrow{W}) \), where \( b, c \in \text{fn}(\Delta_1) \) and \( a \in \text{fn}(V) \)
- \( b(\overrightarrow{V}) \leftrightarrow c(\overrightarrow{W}) \), where \( b, c \in \text{fn}(\Delta_1) \) and \( a \in \text{fn}(W) \)

(since \( a \notin \Delta_1 \), it cannot be the case that \( a \) appears as a buffer endpoint).

Lemma 3.4 tells us that either there exists some \( M' \) such that \( M \rightarrow_M M' \); that \( M \) is a value; or that \( M \) is a communication and concurrency construct. Since \( C \Rightarrow \), we have that \( M \) is unable to reduce (as otherwise \( C \) could reduce by E-LiftM). Since \( a \in \text{fn}(M) \) and \( a \) does not have type 1, it cannot be the case that \( M \) is a value.

Therefore, we have that \( M \) has the form \( E[N] \), where \( N \) is a communication / concurrency construct. This cannot be \textbf{fork}, since \textbf{fork} may always reduce by E-FORK, so there must exist some \( b \in \text{fn}(\Psi, a : S) \) such that ready\((b, M)\).

**Subcase** T-CONNECT\(_2\)

\[
\Psi_1; \Delta_1, a : S \vdash^{\circ} \mathcal{A} \quad \Psi_2, a : S; \Delta_2 \vdash^{\bullet} D
\]

\[
\Psi_1, \Psi_2; \Delta_1, \Delta_2, a : S^\# \vdash^{\bullet} \mathcal{A} \parallel D
\]

\[
\Psi_1, \Psi_2; \Delta_1, \Delta_2 \vdash^{\bullet} (va)(\mathcal{A} \parallel D)
\]
By the definition of auxiliary threads and inversion on the typing relation, we know that $A$ is of the following forms:

- $a(V) \rightsquigarrow b(W)$, where $b \in \Delta_1$
- $b(V) \rightsquigarrow a(W)$, where $b \in \Delta_1$

(as $a \in \text{fn}(A)$ and $a \in \Delta_1$, it cannot be the case that $A$ is a child thread or a zapper thread, as these require empty runtime typing environments).

By the induction hypothesis, we know that $D$ satisfies open progress; hence $(\text{va})(A || D)$ satisfies open progress.

**Case** $C = A || M$

There are three subcases, based on whether the parallel composition arises as a result of $T$-CONNECT$_1$, $T$-CONNECT$_2$, or $T$-MIX.

**Subcase** $T$-CONNECT$_1$

$$\Psi_1, a : S; \Delta_1 \vdash^\circ A \quad \Psi_2, a : S \vdash^* M$$

$$\Psi_1, \Psi_2; \Delta_1, \Delta_2, a : S^\# \vdash^* A || M$$

By inversion on the typing rules, we have that $A$ may be:

- A child thread $\circ M$, where $a \in \text{fn}(M)$
- A zapper thread $\sharp a$
- A buffer $b(V) \rightsquigarrow c(W)$, where $b, c \neq a$ and either $a \in \text{fn}(V)$ or $a \in \text{fn}(W)$

In the case of (1), by Lemma 3.4, we have that either $M$ is a value; there exists $N$ such that $M \rightarrow^M N$; or $M = E[N]$ for some $E, N$, where $N$ is a communication/concurrency construct.

By $T$-CHILD, $\Psi_1, a : S \vdash M : 1$. Since $a \in \text{fn}(M)$ and the only value with type 1 is the unit value $()$ it therefore cannot be the case that $M$ is a value. Since $C \Rightarrow$, it cannot be the case that $M \rightarrow^M N$, since otherwise $C$ could reduce. Thus, it must be the case that $M = E[N]$ where $N$ is a communication and concurrency construct; by similar reasoning as above cases, it therefore must be the case that $\text{ready}(b, M)$ for some $b \in \text{fn}(\Psi_1, a : S)$.

(2) and (3) satisfy the required conditions by definition.

**Subcase** $T$-CONNECT$_2$

$$\Psi_1; \Delta_1, a : S \vdash^\circ A; \Psi_2, a : S; \Delta_2 \vdash^* M$$

$$\Psi_1, \Psi_2; \Delta_1, \Delta_2, a : S^\# \vdash^* A || M$$

Since the runtime typing environment $\Delta_1, a : S$ is non-empty, it cannot be the case that $A$ is a child thread or zapper thread. Thus, $A$ must either be of the form:

1. $a(V) \rightsquigarrow b(W)$, where $a, b \in \Delta_1$; or
2. $b(V) \rightsquigarrow a(W)$, where $a, b \in \Delta_1$

which satisfy the required conditions by definition.

**Subcase** $T$-MIX

$$\Psi_1; \Delta_1 \vdash^\circ A \quad \Psi_2; \Delta_2 \vdash^* M$$

$$\Psi_1, \Psi_2; \Delta_1, \Delta_2 \vdash^* A || M$$

By inversion on the typing rules, we have that $A$ may either be:

1. A child thread $\circ M$
2. A zapper thread $\sharp a$ for some $a \in \text{fn}(\Psi_1)$

(3) A buffer thread $a(V)\rightarrow b(W)$ for some $a, b \in \text{fn}(\Delta_i)$

By Lemma 3.4, we have that $M$ is either a value $V$; there exists some $N$ such that $M \rightarrow_M N$; or $M = E[N]$ for some $E, N$ such that $N$ is a communication and concurrency primitive. It cannot be the case that $M \rightarrow_M N$ since otherwise the configuration could reduce.

By $T$-CHILD, it must be the case that $\Psi_1; \Delta \vdash M : 1$; if $M$ is a value then by inversion on the term typing rules, it must be the case that $M = ()$.

Following the same reasoning as previous cases, if $M = E[N]$ for some communication / concurrency primitive $N$, it must be that $\text{ready}(a, M)$ for some $a \in \Psi_1$.

By the induction hypothesis, we know that $M$ satisfies open progress; hence $\mathcal{A} \parallel M$ satisfies open progress.

**Case $C = \mathcal{T}$**

Assumption: $\Psi; \Delta \vdash^* \mathcal{T}$. By the definition of $\mathcal{T}$, we have two subcases:

**Subcase $\mathcal{T} = \bullet M$**

$\Psi \vdash M : A$  

$$\Psi; \cdot \vdash^* \bullet M$$

By Lemma 3.4, we have that either $M$ is a value; that there exists some $N$ such that $M \rightarrow_M N$; or that there exist some $E, N$ such that $M = E[N]$ where $N$ is a communication / concurrency primitive.

Again, as $\mathcal{C} \Rightarrow$, it cannot be the case that $M \rightarrow_M N$, since otherwise $C$ could reduce. If $M$ is a value, then $\mathcal{T}$ satisfies open progress.

Finally, if $M = E[N]$ where $N$ is a communication / concurrency primitive, it cannot be the case that $N = \text{fork} M'$ since it could reduce by $T$-Fork, and so it must be the case that $\text{ready}(a, M)$ for some $a \in \Psi$, satisfying open progress, as required.

**Subcase $\mathcal{T} = \text{halt}$**

Immediate by the definition of open progress.

$\square$

**Theorem 3.9** provides a more global and concise view of the properties exhibited by a non-reducing process in canonical form, and arises as an immediate corollary.

**Theorem 3.9** Suppose $\Psi; \Delta \vdash^* C$ where $C$ is in canonical form and $C \Rightarrow$.

Let $C = (v_{a_1})(\mathcal{A}_1 \parallel (v_{a_2})(\mathcal{A}_2 \parallel \cdots \parallel (v_{a_n})(\mathcal{A}_n \parallel M))\ldots))$.

Either there exists some $C'$ such that $C \Rightarrow C'$, or:

(1) For $1 \leq i \leq n$, each thread in $\mathcal{A}_i$ is either:

(a) a child thread $\circ M$ for which there exists $a \in \{a_j \mid 1 \leq j \leq i\} \cup \text{fn}(\Psi)$ such that $\text{ready}(a, M)$;

(b) a zapper thread $\downarrow a_i$; or

(c) a buffer.

(2) $M = \mathcal{A}'_1 \parallel \cdots \parallel \mathcal{A}'_m \parallel \mathcal{T}$ such that for $1 \leq j \leq m$:

(a) $\mathcal{A}'_j$ is either:

(i) a child thread $\circ N$ such that $N = ()$ or $\text{ready}(a, N)$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$;

(ii) a zapper thread $\downarrow a$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$; or

(iii) a buffer.

(b) Either $\mathcal{T} = \bullet N$, where $N$ is either a value or $\text{ready}(a, N)$ for some $a \in \{a_i \mid 1 \leq i \leq n\} \cup \text{fn}(\Psi) \cup \text{fn}(\Delta)$; or $\mathcal{T} = \text{halt}$.
C.4 Confluence

**Theorem 3.12 (Diamond Property)** If $\Psi; \Delta \vdash^* C$, and $C \implies D_1$, and $C \implies D_2$, then either $D_1 \equiv D_2$, or there exists some $D_3$ such that $D_2 \implies D_3$ and $D_2 \implies D_3$.

**Proof.** As noted in Section 3.4, $\rightarrow_M$ is deterministic and hence confluent due to the setup of term evaluation contexts, and linearity ensures that endpoints to a buffer may not be shared. Consequently, communication actions on different channels may be performed in any order.

Nevertheless, two critical pairs arise due to asynchrony. The first arises when it is possible to send to or receive from a buffer; there is a choice of whether the send or the receive happens first. Both cases reduce to the same configuration after a single further step.

\[
\mathcal{F}[send\ U\ a] \parallel a(V) \rightsquigarrow b(V \cdot W) \parallel \mathcal{F}'[receive\ b]
\]

\[
\mathcal{F}[a] \parallel a(V) \rightsquigarrow b(V \cdot W \cdot U) \parallel \mathcal{F}'[receive\ b]
\]

The second critical pair arises when sending to a buffer where the peer endpoint has a non-empty buffer and has been cancelled. There is a choice as to whether the value at the head of the queue is cancelled before or after the send takes place. Again, both cases reduce to the same configuration after a single further step.

\[
\mathcal{F}[send\ U\ a] \parallel b \parallel a(V) \rightsquigarrow b(V \cdot W)
\]

\[
\mathcal{F}[a] \parallel b \parallel a(V) \rightsquigarrow b(V \cdot W \cdot U)
\]

\[
\mathcal{F}[send\ U\ a] \parallel b \parallel b(V \cdot W) \parallel \mathcal{F}'[(V, b)]
\]

\[
\mathcal{F}[a] \parallel b \parallel b(V \cdot W \cdot U) \parallel \mathcal{F}'[(V, b)]
\]

\[
\mathcal{F}[send\ U\ a] \parallel b \parallel V \parallel a(V) \rightsquigarrow b(W)
\]

\[
\mathcal{F}[a] \parallel b \parallel V \parallel a(V) \rightsquigarrow b(W \cdot U)
\]

\[
\mathcal{F}[send\ U\ a] \parallel b \parallel V \parallel a(V) \rightsquigarrow b(W) \parallel \mathcal{F}'[(V, b)]
\]

\[
\mathcal{F}[a] \parallel b \parallel V \parallel a(V) \rightsquigarrow b(W \cdot U) \parallel \mathcal{F}'[(V, b)]
\]

\[
\mathcal{F}[a] \parallel b \parallel V \parallel a(V) \rightsquigarrow b(W \cdot U) \parallel \mathcal{F}'[(V, b)]
\]

\[\Box\]
SUPPLEMENT TO SECTION 4.1 (METATHEORY OF EGV WITH ACCESS POINTS)

In this section, we prove that the extension of EGV with access points satisfies preservation.

**Lemma D.1 (Preservation, Access Points (Equivalence)).** If $\Gamma; \Delta \vdash^\phi C$ and $C \equiv D$, then $\Gamma; \Delta \vdash^\phi D$

**Proof.** By induction on the derivation of $C \equiv D$. Rule T-Connect$N$ subsumes T-Connect$_1$ and T-Connect$_2$, so the majority of cases are similar to those we have proven in Lemma C.1. We consider the case for associativity in detail.

**Case** $C \parallel (D \parallel E) \equiv (C \parallel D) \parallel E$

\[
\Gamma = \Gamma_1 + \Gamma''
\]

\[
\Gamma'' = \Gamma_2 + \Gamma_3
\]

\[
\Gamma_2, b_1 : T_1, \ldots, b_m' : T_m, c : S', \Delta_2, a_1 : \overline{S_1}, \ldots, a_m : \overline{S_m}, d : \overline{T'}, \varphi_2 D
\]

\[
\Gamma_3, b_{m+1}', \ldots, b_{n'} : T_n', \varphi_2 D
\]

\[
\Gamma_1, a : S; \Delta_1, b : \overline{T} \varphi_1 C
\]

\[
\Gamma', b : T; \Delta_2, \Delta_3, a : S', c : S'', d : \overline{T'} \varphi_1 \varphi_2 \varphi_3 D \parallel E
\]

\[
\Gamma; \Delta_1, \Delta_2, \Delta_3, a : S', b : T', c : S'', d : \overline{T'} \varphi_1 \varphi_2 \varphi_3 C \parallel (D \parallel E)
\]

where

\[
a : \overline{S} = a_1 : \overline{S}_1, \ldots, a_m : \overline{S}_m, \ldots, a_n : \overline{S}_n
\]

\[
b : \overline{T} = b_1 : \overline{T}_1, \ldots, b_{m'} : \overline{T}_{m'}, \ldots, b_{n'} : \overline{T}_{n'}
\]

\[
\Gamma = \Gamma''' + \Gamma_3
\]

The lemmas for subterm typeability and replacement are slightly different as we must consider unrestricted environments.

**Lemma D.2 (Typeability of Subterms (Access Points)).** If $D$ is a derivation of $\Gamma \vdash E[M] : A$, then there exist $\Gamma_1, \Gamma_2$ and $B$ such that $\Gamma = \Gamma_1 + \Gamma_2$, that $D$ has a subderivation $D'$ that concludes $\Gamma_2 + M : B$, and the position of $D'$ in $D$ corresponds to the position of the hole in $E$.

**Proof.** By induction on the structure of $E$.

**Lemma D.3 (Replacement (Access Points)).** If:

1. $D$ is a derivation of $\Gamma \vdash E[M] : A$, such that $\Gamma = \Gamma_1 + \Gamma_2$
2. $D'$ is a subderivation of $D$ concluding $\Gamma_2 + M : B$
The position of $D'$ in $D$ corresponds to that of the hole in $E$.

Then $\Gamma' \vdash E[N] : A$.

**Proof.** By induction on the structure of $E$. □

**Theorem D.4 (Preservation, Access Points).** If $\Gamma; \Delta \vdash \phi \quad C \implies D$, then $\Gamma; \Delta \vdash \phi \quad D$.

**Proof.** Recall that $\implies$ is defined as $\equiv \rightarrow \equiv$. Therefore, the result arises by induction on the derivation of $C \implies D$ and as a corollary of Lemma D.1.

Again, since T-ConnectN subsumes T-Connect$_1$ and T-Connect$_2$, it suffices only to prove the new cases required for access point reduction.

**Case** E-Spawn

**Assumption:**

$$\Gamma \vdash E[\text{spawn } M] : C$$

$$\Gamma_1; \cdot \vdash^* \bullet E[\text{spawn } M]$$

By Lemma D.2, we have that $\Gamma = \Gamma_1 + \Gamma_2$, and

$$\Gamma_2 \vdash M : 1$$

$$\Gamma_2 \vdash \text{spawn } M : 1$$

By Lemma D.3, we have that $\Gamma_1 \vdash E[()] : C$.

Recomposing:

$$\Gamma = \Gamma_1 + \Gamma_2 \quad \Gamma_1; \cdot \vdash^* E[()] \quad \Gamma_2; \cdot \vdash^* \circ M$$

as required.

**Case** E-New

**Assumption:**

$$\Gamma \vdash E[\text{new}_S] : C$$

$$\Gamma_1; \cdot \vdash^* \bullet E[\text{new}_S]$$

By Lemma D.2 and TA-New, we have that $\cdot \vdash \text{new}_S : AP(S)$.

By Lemma D.3, we have that $\Gamma, z : AP(S) \vdash E[z] : C$.

Thus, we can show:

$$\Gamma, z : AP(S) \vdash E[z] : C$$

$$\Gamma, z : AP(S); \cdot \vdash^* \bullet E[z] \quad ; z : AP(S) \vdash^* z(\epsilon, \epsilon)$$

$$\Gamma, z : AP(S); z : S \vdash^* \bullet E[z] \quad || z(\epsilon, \epsilon)$$

as required.

**Case** E-Accept
Assumption:
\[ \Gamma \vdash \mathbf{E}\{\text{accept } z\} : C \]
\[ \Gamma; \vdash \top \cdot \mathbf{E}\{\text{accept } z\} \quad \vdash z : S, X : \bar{S}, Y : S \vdash^\circ z(X, Y) \]
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ \top \cdot \mathbf{E}\{\text{accept } z\} \parallel z(X, Y) \]

By Lemma D.2, we have that \( \Gamma = \Gamma_1 + \Gamma_2 \) and that \( \Gamma_2 \vdash \text{accept } z : S \). Thus by TA-ACCEPT we have that \( z : \text{AP}(S) \in {}^* \Gamma \).

By Lemma D.3, we have that \( \Gamma, a : S \vdash E[a] : C \).

Recomposing, we have that:
\[ \Gamma, a : S \vdash E[a] : C \]
\[ \Gamma, a : S \vdash \top \cdot \mathbf{E}[a] \]
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ z\{a\} \cup X, Y \]

Recomposing, we have that:
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ \top \cdot \mathbf{E}[a] \parallel z\{a\} \cup X, Y \]

**Case E-Request**

Assumption:
\[ \Gamma \vdash \mathbf{E}\{\text{request } z\} : C \]
\[ \Gamma; \vdash \top \cdot \mathbf{E}\{\text{request } z\} \quad \vdash z : S, X : \bar{S}, Y : S \vdash^\circ z(X, Y) \]
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ \top \cdot \mathbf{E}\{\text{accept } z\} \parallel z(X, Y) \]

By Lemma D.2, we have that \( \Gamma = \Gamma_1 + \Gamma_2 \) and that \( \Gamma_2 \vdash \text{request } z : \bar{S} \). Thus by TA-REQUEST we have that \( z : \text{AP}(S) \in {}^* \Gamma \).

By Lemma D.3, we have that \( \Gamma, a : \bar{S} \vdash E[a] : C \). As duality is involutive, we have that \( \bar{S} = S \).

Recomposing, we have that:
\[ \Gamma, a : \bar{S} \vdash E[a] : C \]
\[ \Gamma, a : \bar{S} \vdash \top \cdot \mathbf{E}[a] \]
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ z(X, \{a\} \cup Y) \]

Recomposing, we have that:
\[ \Gamma ; z : S, X : \bar{S}, Y : S \vdash^\circ \top \cdot \mathbf{E}[a] \parallel z(X, \{a\} \cup Y) \]

as required.

**Case E-Match**

Assumption:
\[ \vdash z : S, a : \bar{S}, X : \bar{S}, b : S, Y : S \vdash^\circ z\{a\} \cup X, \{b\} \cup Y \]

Recomposing:
\[ \bar{S}/e = \bar{S}/e \quad \vdash e : e \quad \vdash e : e \]
\[ \vdash z : S, X : \bar{S}, Y : S \vdash^\circ z(X, Y) \]
\[ \vdash a : \bar{S}, b : S \vdash^\circ a(e) \sim b(e) \]
\[ \vdash z : S, a : \bar{S}, X : \bar{S}, b : S, Y : S \vdash^\circ z(X, Y) \parallel a(e) \sim b(e) \]

\[ \square \]
E DISTRIBUTED DELEGATION

A key feature of \( \pi \)-calculus is mobility, that is, sending channel names as values. In session-based languages and calculi, mobility is realised as session delegation, allowing session-typed channel endpoints to be sent over other session-typed channels. We saw an example of session delegation in §6, in the ChatClient type:

```plaintext
typename ChatClient =!Nickname.
[&|Join:?{Topic, [Nickname], ClientReceive).ClientSend,
Nope:End|&];
```

An endpoint of type ClientReceive is passed as a message.

E.1 Challenges of Distributed Delegation

Session delegation is a vital abstraction in session-based programming. However, its integration with both asynchrony and distribution brings several challenges. The seminal work on distributed delegation is Session Java [Hu et al. 2008].

Fig. 13 shows three scenarios of distributed delegation, as described by Hu et al. [2008]. We write \( X \xrightarrow{x \sim y} Y \) to indicate that \( X \) wishes to send \( x \) to \( Y \) over \( y \) on the basis that \( X \)'s last known location of the corresponding endpoint for \( y \) is \( Y \). Now suppose \( B \xrightarrow{b \sim c} C \). Following Hu et al. [2008], we refer to \( B \) as the session-sender, \( C \) as the session-receiver, and \( A \) as a passive party. There is no happens-before relation between \( A \) sending a message to \( B \) along \( a \), and \( B \) delegating \( b \) to \( C \) along \( c \). Thus, a message could be sent to \( A \) after \( A \) has given up control of \( a \). Following Hu et al. [2008], we call such messages lost messages.
E.2 Approaches to Distributed Delegation

The simplest safe way to implement distributed delegation is to store all buffers on the server, but this requires a blocking remote call for every receive operation. A second naïve method is indefinite redirection, where the session-sender indefinitely forwards all messages to the session receiver. This retains buffer locality, but requires the session-sender to remain online for the duration of the delegated session.

Hu et al. [2008] describe two more realistic distributed delegation algorithms: a resending protocol, which re-sends lost messages after a connection for the delegated session is established, and a forwarding protocol, which forwards lost messages before the delegated session is established. The key idea behind both algorithms is to establish a connection between the passive party and the session-receiver, ensure that the lost messages are received by the session-receiver, and to continue the session only once lost messages are received.

E.3 Delegation in Distributed Session Links

Alas, we cannot directly re-use the resending and forwarding protocols of Hu et al. [2008] because of two fundamental differences in our setting: Links clients do not connect to each other directly, and in Links multiple sessions may be sent at once. Thus, we describe the high-level details of a modified algorithm which addresses these two constraints. We utilise two key ideas:

- Much like the resending protocol, lost messages are retrieved and relayed to the session-receiver once the new session has been established.
- We ensure the session-receiver endpoint is not delegated until the delegation has completed, by queueing messages that include the session-receiver endpoint, and resending them once delegation has completed.

We now consider the case where session-sender and session-receiver are different clients; the case where session-sender is a client and session-receiver the server is similar. Let client A be session-sender and client B be session-receiver.

**Example.** Suppose client A sends a value v containing a session endpoint d along channel (s, t), recalling that s is the peer endpoint and t is the local endpoint. The initial endpoint location table is:

\[ \sigma \triangleq [s \mapsto A, t \mapsto B, b \mapsto A, c \mapsto A] \]
Fig. 14 shows the operation of the delegation protocol on this example. In Step 1, $A$ sends a message to the server $S$, containing the peer endpoint $t$, value to send $v$, and the buffer $\overrightarrow{V}$ for $b$, before beginning to record lost messages for $b$. Upon receiving this message, the server updates its internal mapping for the location of $b$ to be $B$, adds $t$ to the set of delegation carriers $\delta$, and sends a Deliver message containing $t$, $v$, and $\overrightarrow{V}$, before sending a GetLostMessages request to $A$. Upon receiving this message, $A$ will stop recording lost messages for $b$, and relay the lost messages $\overrightarrow{W}$ for $b$ to $S$. The server then sends a Commit message containing $t$ and the lost messages for all delegated endpoints, and removes $t$ from the set of delegation carriers.

The final buffer for $b$ is the concatenation of the initial buffer $\overrightarrow{V}$, the lost messages $\overrightarrow{W}$, and all messages $\overrightarrow{U}$ received for $b$ before the Commit message.

E.4 Correctness

We argue correctness of the algorithm in a similar manner to Hu et al. [2008]. Due to co-operative threading, we can treat each sequence of actions happening at a single participant (for example, steps 3–8) as atomic. Since (as per step 3) the endpoint location table is updated prior to the lost message request, we can safely split the buffer of the delegated session into three parts: the initial buffer being delegated ($\overrightarrow{V}$); the lost messages ($\overrightarrow{W}$); and the messages received after the change in the lookup table but before the Commit message is received ($\overrightarrow{U}$) and reassemble them, retaining ordering.

In our setting, since session channels are not associated with sockets, simultaneous delegation (Fig. 13b) can be handled in the same way as simple delegation. In the case of entangled delegation (Fig. 13c), since delegation carriers may not be delegated themselves until the lost messages have been received, we can be sure that the lost message requests are sent to the correct participant. Hence, the case devolves to simple delegation.