Projet Ajacs

Deliverable WP1

Translation of a high-level language to our sub-language

August 2017

We have chosen ProScript as our security sub-language. Section 4 of the attached paper describes the language.

As ProScript is a typed functional subset of JavaScript, we reuse the compiler developed for JSExplain. See report "Final report on the mechanization of the full JavaScript language" for its description.

Automated Verification for Secure Messaging Protocols and their Implementations: A Symbolic and Computational Approach

Nadim Kobeissi
INRIA Paris

Karthikeyan Bhargavan
INRIA Paris

Bruno Blanchet
INRIA Paris

***Abstract*—Many popular web applications incorporate end-to-end secure messaging protocols, which seek to ensure that messages sent between users are kept confidential and authenticated, even if the web application's servers are broken into or otherwise compelled into releasing all their data. Protocols that promise such strong security guarantees should be held up to rigorous analysis, since protocol flaws and implementations bugs can easily lead to real-world attacks.**

We propose a novel methodology that allows protocol designers, implementers, and security analysts to collaboratively verify a protocol using automated tools. The protocol is implemented in ProScript, a new domain-specific language that is designed for writing cryptographic protocol code that can both be executed within JavaScript programs and automatically translated to a readable model in the applied pi calculus. This model can then be analyzed symbolically using ProVerif to find attacks in a variety of threat models. The model can also be used as the basis of a computational proof using CryptoVerif, which reduces the security of the protocol to standard cryptographic assumptions. If ProVerif finds an attack, or if the CryptoVerif proof reveals a weakness, the protocol designer modifies the ProScript protocol code and regenerates the model to enable a new analysis.

We demonstrate our methodology by implementing and analyzing a variant of the popular Signal Protocol with only minor differences. We use ProVerif and CryptoVerif to find new and previously-known weaknesses in the protocol and suggest practical countermeasures. Our ProScript protocol code is incorporated within the current release of Cryptocat, a desktop secure messenger application written in JavaScript. Our results indicate that, with disciplined programming and some verification expertise, the systematic analysis of complex cryptographic web applications is now becoming practical.

1. Introduction

Designing new cryptographic protocols is highly error-prone; even well-studied protocols, such as Transport Layer Security (TLS), have been shown to contain serious protocol flaws found years after their deployment (see e.g. [1]). Despite these dangers, modern web applications often embed custom cryptographic protocols that evolve with each release. The design goal of these protocols is typically to protect user data as it is exchanged over the web and synchronised across devices, while optimizing performance for application-specific messaging patterns and deployment constraints. Such custom protocols deserve close scrutiny, but their formal security analysis faces several challenges.

First, web applications often evolve incrementally in an ad hoc manner, so the embedded cryptographic protocol is only ever fully documented in source code. Even when protocol designers or security researchers take the time to create a clear specification or formal model for the protocol, these documents are typically incomplete and quickly go out-of-date. Finally, even if the protocol itself is proved to be secure, bugs in its implementation can often bypass the intended security guarantees. Hence, it is not only important to extract a model of the protocol from the source code and analyze its security, it is essential to do so in a way that the model can evolve as the application is modified.

In this paper, we study the protocol underlying the Signal messaging application developed by Open Whisper Systems. Variants of this protocol have also been deployed within WhatsApp, Facebook Messenger, Viber, and many other popular applications, reaching over a billion devices. The protocol has been known by other names in the past, including Axolotl, TextSecure (versions 1, 2, and 3), and it continues to evolve within the Signal application under the name Signal Protocol. Until recently, the main documentation for the protocol was its source code, but new specifications for key components of the protocol have now been publicly released.[1]

Signal Protocol has ambitious security goals; it enables asynchronous (zero round-trip) authenticated messaging between users with end-to-end confidentiality. Each message is kept secret even if the messaging server is compromised, and even if the user's long term keys are compromised, as long as these keys are not used by the attacker before the target message is sent (forward and future secrecy.) To achieve these goals, Signal uses a novel authenticated key exchange protocol (based on mixing multiple Diffie-Hellman shared secrets) and a key refresh mechanism (called double ratcheting). The design of these core mechanisms in TextSecure version 2 was cryptographically analyzed in [2] but the protocol has evolved since then and the security of Signal as it is currently implemented and deployed remains an open question.

In fact, although they all implement the same core protocol, different implementations of the Signal protocol

1. https://whispersystems.org/docs/specifications/x3dh/

vary in important details, such as how users are identified and authenticated, how messages are synchronised across devices, etc. We seek to develop and analyze one such variant that was recently incorporated into Cryptocat, a desktop messaging application developed by one of the current authors. We call this variant SP in the rest of this paper. We develop a detailed model for SP in the applied pi calculus and verify it using the ProVerif protocol analyzer [3] for these security goals against adversaries in a classic Dolev-Yao model [4]. We also develop a computational proof for SP using the CryptoVerif prover [5]. There remains the challenge of keeping our models up-to-date as the protocol code evolves within Cryptocat. To this end, we design a model extraction tool that can compile the protocol source code to the applied pi calculus.

Signal has been implemented in various programming languages, but most desktop implementations of Signal, including Cryptocat, are written in JavaScript. Although JavaScript is convenient for widespread deployability, it is not an ideal language for writing security-critical applications. Its permissive, loose typing allows for dangerous implementation bugs and provides little isolation between verified cryptographic protocol code and unverified third-party components. Rather than trying to verify general JavaScript programs, we advocate that security-critical components like SP should be written in a well-behaved subset that enables formal analysis.

We introduce ProScript (short for *"Protocol Script"*), a programming and verification framework tailored specifically for the implementation of cryptographic protocols. ProScript extends Defensive JavaScript (DJS) [6], [7], a static type system for JavaScript that was originally developed to protect security-critical code against untrusted code running in the same origin. ProScript is syntactically a subset of JavaScript, but it imposes a strong coding discipline that ensures that the resulting code is amenable to formal analysis. ProScript programs are mostly self-contained; they cannot call arbitrary third-party libraries, but are given access to carefully implemented (well-typed) libraries such as the ProScript cryptographic library (PSCL). Protocols written in ProScript can be type-checked and then automatically translated into an applied pi calculus [8] model using the ProScript compiler. The resulting model can be analyzed directly through ProVerif and can be adapted and extended to a proof in CryptoVerif. As the code evolves, this model can be automatically refreshed to enable new analyses and proofs, if necessary.

Contributions. We present an outline of our contributions in this paper:

A Security Model and New Attacks. We present security goals and a threat model for secure messaging (§ 2). As a motivation for our verification approach, we discuss protocol weaknesses and implementation bugs in the messaging protocol underlying the popular Telegram application.

Automated Model Extraction from JavaScript. We present the ProScript compiler, which allows for the compilation from a subset of JavaScript into a readable protocol model in the applied pi calculus (§4). Model extraction enables formal verification to keep up with rapidly changing source code. Readable models allow the protocol analyst to experiment with different threat models and security goals and to test new features before including them in the implementation.

A Symbolic Security Analysis of SP. We formalize and analyze a variant of Signal Protocol for a series of security goals, including confidentiality, authenticity, forward secrecy and future secrecy, against a classic symbolic adversary (§5). Our analysis uncovers several weaknesses, including previously unreported replay and key compromise impersonation attacks, and we propose and implement fixes which we then also verify.

A Computational Cryptographic Proof for SP. We present proofs of message authenticity, secrecy and forward secrecy for SP obtained using the CryptoVerif computational model prover [5]. (§6)

A Verified Protocol Core for Cryptocat. We integrate our verified protocol code into the latest version of Cryptocat[2] (§7), a popular open source messaging client with thousands of users that is developed and maintained by one of the authors of this paper. We show how the new architecture of Cryptocat serves to protect the verified protocol code from bugs in the rest of the application.

2. A Security Model for Encrypted Messaging

We consider a simple messaging API as depicted below. An initiator A can start a conversation with B by calling `startSession` with long-term secrets for A and any identity credentials it has for B. This function returns the initial conversation state T_0. Thereafter, the initiator can call `send` with a plaintext message M_1 to obtain the encrypted message E_1 that it needs to send on the network. Or it can call `recv` with an encrypted message E_2 it received (supposedly from B) to obtain the plaintext message M_2.

```
T_0^ab = startSession(secrets_A, identity_B)
T_1^ab, E_1 = send(T_0^ab, M_1)
T_2^ab, M_2 = recv(T_1^ab, E_2)
```

The responder B uses a similar API to accept sessions and receive and send messages:

```
T_0^ba = acceptSession(secrets_B, identity_A)
T_1^ba, M_1 = recv(T_0^ba, E_1)
T_2^ba, E_2 = send(T_1^ba, M_2)
```

We deliberately chose a functional state-passing API with no side-effects in order to focus on cryptographic protocol computations, rather than the concrete details of how these messages are sent over the network.

2.1. Threat Model

While threat models vary for different protocols, we consider the following threats in this paper:

2. https://crypto.cat

- **Untrusted Network** We assume that the attacker controls the network and so can intercept, tamper with and inject network messages (e.g. E_1, E_2). Moreover, if two messaging clients communicate via a server, we typically treat that server as untrusted.
- **Malicious Principals** The attacker controls a set of valid protocol participants (e.g. M), for whom it knows the long-term secrets. The attacker may advertise any identity key for its controlled principals; it may even pretend to own someone else's identity keys.
- **Long-term Key Compromise** The attacker may compromise a particular principal (e.g. A) during or after the protocol, to obtain her long-term secrets.
- **Session State Compromise** The attacker may compromise a principal to obtain the full session state at some intermediate stage of the protocol (e.g. T_1^{ab}).

2.2. Cryptographic Models

Traditionally, symbolic cryptographic models have been particularly suitable for automated protocol analysis. They ignore attacks with negligible probability and assume that each cryptographic function is a perfect black-box. For example, in such models, hash functions never collide and encryption is a message constructor that can only be reversed by decryption. In the *computational model*, cryptographic primitives are functions over bitstrings and their security is specified in terms of probabilities. These models are more precise and closer to those used by cryptographers, but usually do not lend themselves to fully automated proofs. Generally, we will use symbolic models when we are trying to find attacks that rely on logical flaws in the protocol and in its use of cryptographic primitives. We will use computational models when we want to build a cryptographic proof of security, starting from standard cryptographic assumptions.

2.3. Security Goals

We state a series of semi-formal security goals in terms of our generic messaging API. We use the phrase "A sends a message M to B" to mean that A calls $\mathsf{Send}(T, M)$ with a session state T that represents a conversation between A and B. Similarly, we say that "B receives a message M from A" to mean that B obtained M as a result of calling $\mathsf{Recv}(T, E)$ with a session T with A.

Unless otherwise specified, the following security properties assume that both A and B are honest, that is, their long-term secrets have not been compromised. We begin with several variants of authenticity goals:

- **Message Authenticity** If B receives a message M from A, then A must have sent M to B.
- **No Replays** Each message received by B from A corresponds to a unique message sent by A. That is, the attacker must not be able to get a single message sent by A to be accepted twice at B.
- **No Key Compromise Impersonation** Even if the long-term secrets of B are compromised, message authenticity must hold at B. That is, the attacker must not be able to forge a message from A to B.

Our definition of message authenticity covers integrity as well as sender and recipient authentication. Obtaining message authenticity also helps prevent *unknown key share* attacks, where B receives a message M from A, but A sent that message to a different intended recipient C. We define four confidentiality goals:

- **Secrecy** If A sends some secret message M to B, then nobody except A and B can obtain M.
- **Indistinguishability** If A randomly chooses between two messages M_0, M_1 (of the same size) and sends one of them to B, the attacker cannot distinguish (within the constraints of the cryptographic model) which message was sent.
- **Forward Secrecy** If A sends a secret message M to B and if A's and B's long-term secrets are subsequently compromised, the message M remains secret.
- **Future Secrecy** Suppose A sends M in a session state T, then receives N, then sends M'. If the session state T is subsequently compromised, the message M' remains secret.

Some protocols satisfy a weaker notion of forward secrecy, sometimes called *weak* forward secrecy, where an attacker is not allowed to actively tamper with the protocol until they have compromised the long-term keys [9]. Some messaging protocols also seek to satisfy more specific authenticity and confidentiality goals, such as non-repudiation and plausible deniability. We will ignore them in this paper.

In the next section, we evaluate two secure messaging applications against these goals, we find that they fail some of these goals due to subtle implementation bugs and protocol flaws. Hence, we advocate the use of automated verification tools to find such attacks and to prevent their occurrence.

3. Analyzing Real-World Messaging Protocols

Modern messaging and transport protocols share several distinctive features [10]: for example, Signal Protocol, SCIMP, QUIC and TLS 1.3 share a strong focus on asynchronous key agreement with a minimum of round trips. Some also guarantee new security goals such as future secrecy. The protocols also assume non-standard (but arguably more user-friendly) authentication infrastructures such as Trust-on-First-Use (TOFU). Modern messaging protocols have several interesting features and properties that set them apart from classic cryptographic protocols and upon which we focus our formal verification efforts:

New Messaging Patterns. In contrast to connection-oriented protocols, modern cryptographic protocols are constrained by new communication flows such as zero-round-trip connections and asynchronous messaging, where the peer may not even be present.

Confidentiality Against Strong Adversaries. Web messaging protocols need to be robust against server compromise

and device theft and so seek to provide strong and novel forward secrecy guarantees.

Weak Authentication Frameworks. Many of these protocols do not rely on public key infrastructures. Instead they may authenticate peers on a TOFU basis or even let peers remain anonymous, authenticating only the shared connection parameters.

Code First, Specify Later. Unlike Internet protocols, which are designed in committee, these protocols are first deployed in code and hand-tuned for performance on a particular platform. The code often remains the definitive protocol specification.

Before outlining our verification approach for such protocols, we take a closer look at two messaging applications: Telegram and Cryptocat.

3.1. Secret Chats in Telegram

Our first example is the "MTProto" [11] secure messaging protocol used in the Telegram messaging application. We focus on the "secret chat" feature that allows two Telegram clients who have already authenticated themselves to the server to start an encrypted conversation with each other. Although all messages pass through the Telegram server, the server is untrusted and should not be able to decrypt these messages. The two clients A and B download Diffie-Hellman parameters from the Telegram server and then generate and send their public values to each other.

The key exchange is not authenticated with long-term credentials. Instead, the two clients are expected to communicate out-of-band and compare a SHA-1 hash (truncated to 128-bits) of the Diffie-Hellman shared secret. If two users perform this authentication step, the protocol promises that messages between them are authentic, confidential, and forward secret, even if the Telegram server is compromised. However this guarantee crucially relies on several cryptographic assumptions, which may be broken either due to implementation bugs or computationally powerful adversaries, as we describe below.

Malicious Primes. MTProto relies on clients checking that the Diffie-Hellman configuration (p, g) that they received from the server is suitable for cryptographic use. The specification requires that p be a large safe prime; hence the client must check that it has exactly 2048 bits and that both p and $(p-1)/2$ are prime, using 15 rounds of the Miller-Rabin primality test. There are several problems with this check. First, the server may be able to carefully craft a non-prime that passes 15 rounds of Miller-Rabin. Second, checking primality is not enough to guarantee that the discrete log problem will be hard. If the prime is chosen such that it has "low weight", the SNFS algorithm applies, making discrete logs significantly more practical [12]. Even if we accept that primality checking may be adequate, it is unnecessary for an application like Telegram, which could simply mandate the use of well-known large primes instead [13].

Public Values in Small Subgroups. A man-in-the-middle can send to both A and B public Diffie-Hellman values g^b and g^a equal to 1 (resp. 0, resp. $p-1$). Both A and B would then compute the shared secret as $g^{ab} = 1$ (resp. 0, resp. 1 or -1). Since their key hashes match, A and B think they have a confidential channel. However, the attacker can read and tamper with all of their messages. More generally, MTProto relies on both peers verifying that the received Diffie-Hellman public values do not fall in small subgroups. This check is adequate to prevent the above attack but could be made unnecessary if the two public values were to be authenticated along with the shared secret in the hash compared by the two peers.

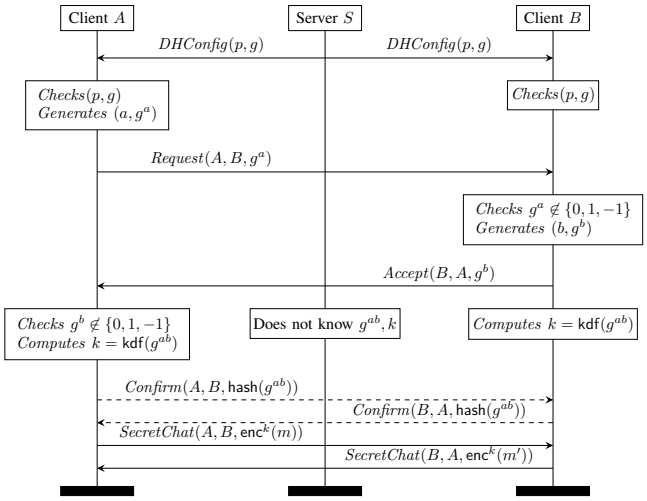

Figure 1: Telegram's MTProto Protocol for Secret Chats.

Implementation Bugs in Telegram for Windows. The above two weaknesses, reported for the first time in this paper, can result in attacks if the protocol is not implemented correctly. We inspected the source code for Telegram on different platforms; while most versions perform the required checks, we found that the source code for Telegram for Windows Phone did not check the size of the received prime, nor did it validate the received Diffie-Hellman values against 1, 0 or $p-1$. We reported both bugs to the developers, who acknowledged them and awarded us a bug bounty.

Such bugs and their consequent attacks are due to missed security-relevant checks, and they can be found automatically by symbolic analysis. For example, [14] shows how to model unsafe (malicious) primes and invalid public keys in ProVerif and uses this model to find vulnerabilities in several protocols that fail to validate Diffie-Hellman groups or public keys.

Other Cryptographic Weaknesses. MTProto is also known to be vulnerable to an authentication attack if an adversary can compute 2^{64} SHA-1 hashes [15], and to chosen-ciphertext attacks on its unusual AES-IGE encryption scheme [16]. How can we be sure that there are no other protocol flaws or implementation bugs hiding in MTProto? Any such guarantee would require a systematic security analysis of both the protocol and the source code against

both symbolic and computational adversaries.

3.2. A New Protocol for Cryptocat

Cryptocat is a secure messaging application that is written in JavaScript and deployed as a desktop web application. Earlier versions of Cryptocat implement a variant of the OTR (Off-The-Record) messaging protocol [17] which suffers from several shortcomings. It does not support asynchronous messaging, so both peers have to be online to be able to message each other. It does not support multiple devices or encrypted file transfer. OTR also uses legacy cryptographic constructions like DSA signatures and prime-field Diffie-Hellman, which are slower and less secure than more modern alternatives based on elliptic curves. Furthermore, Cryptocat peers did not have long-term identities and so the authentication guarantees are weak. Early version of Cryptocat suffered from many high-profile implementation bugs, including the reuse of initialization vectors for file encryption [18], bad random number generation, and a classic JavaScript type flaw that resulted in a private key of 255 bits being coerced into a string that held only 55 bits. Some of these implementation flaws would have been found using a static type checker, others required deeper analysis.

Cryptocat was recently rewritten from scratch to upgrade both its messaging protocol and its implementation. The goal of this redesign was to isolate its protocol core and replace it with a verified messaging protocol written in a statically typed subset of JavaScript.

3.3. Towards Automated Verification

The innovative designs and unusual security guarantees of secure messaging protocols demand formal security analysis. Hand-written models with detailed cryptographic proofs can be useful as a reference, but we observe that the most recent analysis of Signal Protocol [2] is already out of date, as the protocols have moved on to new versions. Furthermore, manual cryptographic proofs often leave out details of the protocol for simplicity and some of these details (e.g. client authentication) may lead to new attacks. In this paper, we advocate the use of automated verification tools to enable the analysis of complex protocols as they evolve and incorporate new features. Moreover, we would also like to find protocol implementation bugs (like the ones in previous versions of Telegram and Cryptocat) automatically.

We advocate the verification approach depicted in Figure 2. The messaging application is written in JavaScript and is broken down into a cryptographic protocol core and untrusted application code that interact through a small well-typed API that hides all protocol secrets within the protocol core and only offers a simple send/receive functionality to the application. Notably, the protocol core is written in a domain-specific language and does not rely on any external libraries except for a well-vetted cryptographic library. The protocol code can be translated to an applied pi calculus

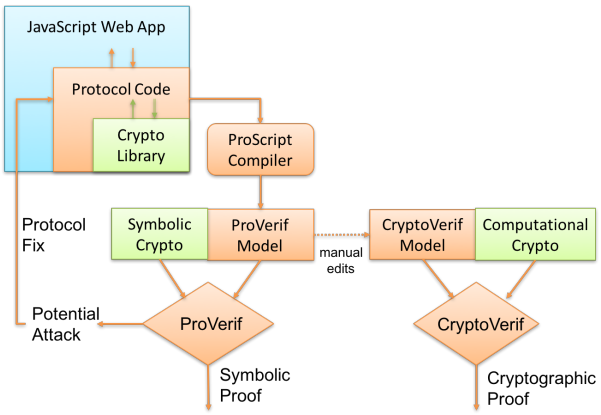

Figure 2: Verification Approach. A ProVerif model is automatically extracted from ProScript protocol code and analyzed for its security goals against a symbolic attacker. The model is then edited by hand and extended with cryptographic assumptions and intermediate lemmas to build a computational proof that is verified by CryptoVerif.

model and symbolically analyzed in ProVerif to find protocol flaws and attacks. The model can also be used as the starting point for a cryptographic proof for the protocol developed using CryptoVerif.

In the rest of this paper, we show how we applied this verification methodology to systematically analyze a variant of the Signal protocol, called SP, that is implemented in the new version of Cryptocat.

4. ProScript: A Language for Protocol Implementation

ProScript aims to be an ideal language for reliably implementing cryptographic protocols for web applications. Using ProScript, a protocol designer or implementer can implement a protocol, automatically extract a formal model from the code, verify the model using ProVerif, and then run the protocol code within a JavaScript web application. The ProScript framework does not target general JavaScript code, however existing applications can be adapted to use ProScript for their security-critical protocol components.

Our goal is to allow the developer to go back and forth between their protocol implementation and the ProVerif model, in order to help understand the behavior being illustrated, the properties being verified and how detected attacks, if any, relate to their source code. For these reasons, we pay special attention to generating models that are optimized both for verifiability as well as readability. This increases their utility to a human examiner who may wish to independently expand the model to include more specific process flows or to explore variations of the protocol against a manually defined adversary.

Syntactically, ProScript is a subset of JavaScript that can be naturally translated to the applied pi calculus. This restriction produces casualties, including recursion, `for`

loops and extensible objects. A closer look at the ProScript syntax shows JavaScript employed in a particular style to bring out useful features:

Isolation. ProScript is based on Defensive JavaScript (DJS) [6], [7], a typed subset of JavaScript which focuses on protecting security-critical components from malicious JavaScript code running in the same environment. DJS imposes a strict typing discipline in order to eliminate language-based attacks like prototype poisoning. In particular, it forbids the use of unknown external libraries as well as calls to tamperable object methods such as `.toString()`. It also forbids extensible objects and arrays and prevents any access to object prototypes. These restrictions result in protocol implementations that are more robust and less influenced by the runtime environment. The ProScript typechecker builds on and extends DJS and hence, inherits both its language restrictions and isolation guarantees.

Type Declarations and Inference. ProScript requires all variables and functions to be declared before they are used, hence imposing a strict scoping discipline. For example, an expression $v.x$ is well-typed if and only if v has been defined, as a local or global variable, to be an object with a property x. As an extension to the builtin types of DJS, ProScript allows type declarations for commonly used protocol data structures. For example, an array of 32 hexadecimal integers can be declared as a key type. The ProScript compiler recognizes such type declarations and uses them to translate the code into more concise and informative ProVerif models. Moreover, the typechecker can automatically infer fine-grained sub-types. For example, ProScript differentiates between numbers declared using decimal literals (ex. `128`) and hexadecimal literals (ex. `0x80`). Numbers defined using hexadecimal are sub-typed as bytes. This feature allows us to track how numerical values are employed in the protocol, and prevents type coercion bugs similar to an actual bug that we describe in §3.2, where a significant loss of entropy was caused by a byte being coerced into a decimal value.

State-Passing Functional Style. ProScript's syntax takes advantage of JavaScript's functional programming features in order to encourage and facilitate purely functional protocol descriptions, which the compiler can translate into symbolically verifiable, human-readable models in the applied pi calculus. The functional style encourages the construction of state-passing functions, leaving state modification up to the unverified application outside of the ProScript code. The majority of a ProScript implementation tends to be a series of pure function declarations. A small subset of these functions is exposed to a global namespace for access by the verified application while most remain hidden as utility functions for purposes such as key derivation, decryption and so on. This state-passing style is in contrast to DJS that allows direct modification of heap data structures. The functional style of ProScript allows protocol data structures, including objects and arrays, to be translated to simple terms in ProVerif built using constructors and destructors, hence avoiding the state-space explosion inherent in the heap-based approach that is needed to translate DJS to ProVerif [7].

4.1. ProScript Syntax

A ProScript implementation consists of a series of *modules*, each containing a sequence of type declarations (containing constructors, assertion utilities, and type converters), constant declarations and function declarations.

Type Declarations. Types are declared as object constants with the name Type_x where x is the name of the type (e.g. key). Type declaration objects include various properties:

1) $construct$, a function which returns the shape of the object and which is used to both define the type to the compiler and to instantiate new variables of this type throughout the code.
2) $assert$, a function allowing the ProScript compiler to infer that a function's input parameter is of type x.
3) $toBitstring, fromBitstring$, conversion functions from type x to a regular string. These are detected by the ProScript compiler and represented as ProVerif type conversion functions in the extracted model.
4) $clone$, a function for cloning a variable of type x.

Constant Declarations. In ProScript, we prefer the use of constant declarations to variable declarations due to their decreased malleability. We also prohibit the reassignment of object properties because of the pointer-like behavior of JavaScript object variables rendering this difficult to model efficiently in ProVerif. Scoping is also enforced: ProScript only allows for the declaration of variables at the top of a module or function, preceding all function declarations or calls. Further, for an object v, we say that $v.x$ is well-formed if and only if v has been declared, either as a type or scoped variable, to have a property x.

Function Declarations. The majority of a ProScript implementation tends to be a series of pure function declarations. A small subset of these functions is exposed to a global namespace for access by the verified application while most remain hidden as utility functions for purposes such as key derivation, decryption and so on. All ProScript functions are pure and state-passing, allowing them to be modeled into ProVerif functions, declared with the keyword *letfun* (See §4.2). All ProScript functions are typed and the ProScript compiler will enforce that they are used with the same input types and return the same output type throughout the implementation. Top-level object declarations often contain function declarations. We assume for simplicity that these object declarations are flattened into top-level function declarations. In order to enforce type safety, ProScript objects and arrays are also non-extensible: JavaScript prototype functions such as `Array.push` cannot be employed and object accessors such as $v[x]$ are disallowed.

ProScript

$v ::=$	values
x	variables
n	numbers
s	strings
`true, false`	booleans
`undefined, null`	predefined constants

$e ::=$	expressions
v	values
$\{x_1 : v_1, \ldots, x_n : v_n\}$	object literals
$v.x$	field access
$[v_1, \ldots, v_n]$	array literals
$v[n]$	array access
$\texttt{Lib}.l(v_1, \ldots, v_n)$	library call
$f(v_1, \ldots, v_n)$	function call
$\sigma ::=$	statements
$\texttt{var}\ x; \sigma$	variable declaration
$x = e; \sigma$	variable assignment
$\texttt{const}\ x = e; \sigma$	constant declaration
$\texttt{if}\ (v_1 === v_2)\ \{\sigma_1\}\ \texttt{else}\ \{\sigma_2\}$	if-then-else
$\texttt{return}\ e$	return
$\gamma ::=$	globals
$\texttt{const}\ x = e$	constants
$\texttt{const}\ f = \texttt{function}(x_1, \ldots, x_n)\{\sigma\}$	functions
$\texttt{const Type\_}x = \{\ldots\}$	user types
$\mu ::= \gamma_0; \ldots; \gamma_n$	modules

Note that we will use the defined Lib.l notation to access the ProScript Cryptography Library.

Operational Semantics. ProScript's operational semantics is a subset of JavaScript, and both run on JavaScript interpreters. It is tooled based on the formal semantics of Maffeis et al. [19] and is superficially adapted for our language subset.

4.2. ProVerif Syntax

A ProVerif script Σ is divided into two major parts:

1) $\Delta_1 \ldots \Delta_n$, a sequence of declarations which encapsulates all types, free names, queries, constructors, destructors, equations, pure functions and processes. Queries define the security properties to prove. Destructors and equations define the properties of cryptographic primitives.
2) P, the top-level process which then effectively employs $\Delta_1 \ldots \Delta_n$ as its toolkit for constructing a process flow for the protocol.

In processes, the replication $!P$ represents an unbounded number of copies of P in parallel. Tables store persistent state: The process $\texttt{insert}\ a(M_1, \ldots, M_n); P$ inserts the entry $(M_1, \ldots, M_n)$ in table a, and runs P. The process $\texttt{get}\ a(=M_1, x_2, \ldots, x_n)\ \texttt{in}\ P$ looks for an entry $(N_1, \ldots, N_n)$ in table a such that $N_1 = M_1$. When such an entry is found, it binds $x_2, \ldots, x_n$ to $N_2, \ldots, N_n$ respectively and runs P. Events are used for recording that certain actions happen (e.g. a message was sent or received), in order to use that information for defining security properties. Phases model a global synchronization: processes initially run in phase 0; then at some point processes of phase 0 stop and processes of phase 1 run and so on. For instance, the protocol may run in phase 0 and some keys may be compromised after the protocol run by giving them to the adversary in phase 1.

ProVerif

$M ::=$	terms
v	values
a	names
$f(M_1, \ldots, M_n)$	function application
$E ::=$	enriched terms
M	return value
$\texttt{new}\ a : \tau; E$	new name a of type τ
$\texttt{let}\ x = M\ \texttt{in}\ E$	variable definition
$\texttt{if}\ M = N\ \texttt{then}\ E_1\ \texttt{else}\ E_2$	if-then-else
$P, Q ::=$	processes
0	null process
$\texttt{in}(M, x : \tau); P$	input x from channel M
$\texttt{out}(M, N); P$	output N on channel M
$\texttt{let}\ x = M\ \texttt{in}\ P$	variable definition
$P \mid Q$	parallel composition
$!P$	replication of P
$\texttt{insert}\ a(M_1, \ldots, M_n); P$	insert into table a
$\texttt{get}\ a(=M_1, x_2, \ldots, x_n)\ \texttt{in}\ P$	get table entry specified by M_1
$\texttt{event}\ M; P$	event M
$\texttt{phase}\ n; P$	enter phase n
$\Delta ::=$	declaration
$\texttt{type}\ \tau$	type τ
$\texttt{free}\ a : \tau$	name a
$\texttt{query}\ q$	query q
$\texttt{table}\ a(\tau_1, \ldots, \tau_n)$	table a
$\texttt{fun}\ C(\tau_1, \ldots, \tau_n) : \tau$	constructor
$\texttt{reduc forall}\ x_1 : \tau_1, \ldots, x_n : \tau_n; f(M_1, \ldots, M_n) = M$	destructor
$\texttt{equation forall}\ x_1 : \tau_1, \ldots, x_n : \tau_n; M = M'$	equation
$\texttt{letfun}\ f(x_1 : \tau_1, \ldots, x_n : \tau_n) = E$	pure function
$\texttt{let}\ p(x_1 : \tau_1, \ldots, x_n : \tau_n) = P$	process
$\Sigma ::= \Delta_1 \ldots \Delta_n.\texttt{process}\ P$	script

4.3. Translation

Within Σ, ProScript functions are translated into ProVerif pure functions. Type declarations are translated into ProVerif type declarations. Individual values, such as strings and numbers, are declared as global constants at the top-level scope of the ProVerif model with identifiers that are then employed throughout the model when appropriate. Objects and Arrays are instantiated in the model using functions, with destructors automatically generated in order to act as getters.

Translation Rules

$M_v ::= v \mid \{x_1 : v_1, \ldots, x_n : v_n\} \mid [v_1, \ldots, v_n]$	
$\mathcal{V}[\![M_v]\!] \to M$	Values to Terms
$\mathcal{V}[\![v]\!] = v$	
$\mathcal{V}[\![\{x_1 : v_1, \ldots, x_n : v_n\}]\!] = \texttt{Obj\_t}(v_1, \ldots, v_n)$	
$\mathcal{V}[\![[v_1, \ldots, v_n]]\!] = \texttt{Arr\_t}(v_1, \ldots, v_n)$	
$\mathcal{E}[\![e]\!] \to M$	Expressions to Terms

$\mathcal{E}[\![M_v]\!] = \mathcal{V}[\![M_v]\!]$
$\mathcal{E}[\![v.x]\!] = \texttt{get\_x}(v)$
$\mathcal{E}[\![v[i]]\!] = \texttt{get\_i}(v)$
$\mathcal{E}[\![\texttt{Lib.}l(v_1, \ldots, v_n)]\!] = \texttt{Lib\_}l(\mathcal{V}[\![v_1]\!], \ldots, \mathcal{V}[\![v_n]\!])$
$\mathcal{E}[\![f(v_1, \ldots, v_n)]\!] = f(\mathcal{V}[\![v_1]\!], \ldots, \mathcal{V}[\![v_n]\!])$

$\mathcal{S}[\![\sigma]\!] \to E$ Statements to Enriched Terms
$\mathcal{S}[\![\texttt{var } x; \sigma]\!] = \mathcal{S}[\![\sigma]\!]$
$\mathcal{S}[\![x = e; \sigma]\!] = \texttt{let } x = \mathcal{E}[\![e]\!] \texttt{ in } \mathcal{S}[\![\sigma]\!]$
$\mathcal{S}[\![\texttt{const } x = e; \sigma]\!] = \texttt{let } x = \mathcal{E}[\![e]\!] \texttt{ in } \mathcal{S}[\![\sigma]\!]$
$\mathcal{S}[\![\texttt{return } v]\!] = \mathcal{V}[\![v]\!]$
$\mathcal{S}[\![\texttt{if } (v_1 \texttt{===} v_2) \{\sigma_1\} \texttt{ else } \{\sigma_2\}]\!] =$
$\quad \texttt{if } \mathcal{V}[\![v_1]\!] = \mathcal{V}[\![v_2]\!] \texttt{ then } \mathcal{S}[\![\sigma_1]\!] \texttt{ else } \mathcal{S}[\![\sigma_2]\!]$

$\mathcal{F}[\![\gamma]\!] \to \Delta$ Types and Functions to Declarations
$\mathcal{F}[\![\texttt{const } f = \texttt{function}(x_1, \ldots, x_n)\{\sigma\}]\!] =$
$\quad \texttt{letfun } f(x_1, \ldots, x_n) = \mathcal{S}[\![\sigma]\!]$
$\mathcal{F}[\![\texttt{const Type\_}t = \{\ldots\}]\!] = \texttt{type } t$

$\mathcal{C}[\![\mu]\!](P) \to P$ Constants to Top-level Process
$\mathcal{C}[\![\epsilon]\!](P) = P$
$\mathcal{C}[\![\texttt{const } x = e; \mu]\!](P) = \texttt{let } x = \mathcal{E}[\![e]\!] \texttt{ in } \mathcal{C}[\![\mu]\!](P)$

$\mathcal{M}[\![\mu]\!](P) \to \Sigma$ Modules to Scripts
$\mathcal{M}[\![\mu]\!](P) = \mathcal{F}[\![\gamma_1]\!]. \ldots \mathcal{F}[\![\gamma_n]\!].\mathcal{C}[\![\mu_c]\!](P)$
where μ_c contains all globals $\texttt{const } x = e$ in μ
and $\gamma_1, \ldots, \gamma_n$ are the other globals of μ.

Translation Soundness. We currently do not formally prove translation soundness, so proofs of the resulting ProVerif model do not necessarily imply proof of the source code. Instead, we use model translation as a pragmatic tool to automatically generate readable protocol models faithful to the implementation, and to find bugs in the implementation. We have experimented with multiple protocols written in ProScript, including OTR, SP, and TLS 1.3, and by carefully inspecting the source code and target models, we find that the compiler is quite reliable and that it generates models that are not so far from what one would want to write directly in ProVerif. In future work, we plan to prove the soundness of this translation to get stronger positive guarantees from the verification. To this end, we observe that our source language is a simply-typed functional programming language, and hence we should be able to closely follow the methodology of [20].

Generating Top-Level Processes. We are also able to automatically generate top-level ProVerif processes. Aiming to implement this in a way that allows us to easily integrate ProScript code into existing codebases, we decided to describe top-level functions inside `module.exports`, the export namespace used by modules for Node.js [21], a popular client/server run-time for JavaScript applications (based on the V8 engine). This makes intuitive sense: `module.exports` is used specifically in order to define the functions of a Node.js module that should be available to the external namespace once that module is loaded, and executing all this functionality in parallel can give us a reasonable model of a potential attacker process. Therefore, functions declared in this namespace will be translated into top-level processes executed in parallel. We use ProVerif `tables` in order to manage persistent state between these parallel processes: each process fetches the current state from a table, runs a top-level function in that state, and stores the updated state returned by the function in the table.

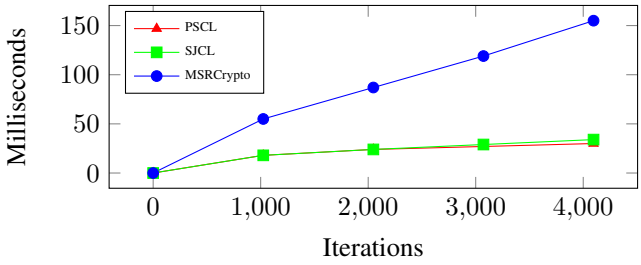

Figure 3: Each SHA256 iteration hashes 16 blocks.

4.4. Trusted Libraries for ProScript

Protocol implementations in ProScript rely on a few trusted libraries, in particular, for cryptographic primitives and for encoding and decoding protocol messages.

When deployed in Node.js or within a browser, the protocol code may have access to native cryptographic APIs. However, these APIs do not typically provide all modern cryptographic primitives; for example, the W3C Web Cryptography API does not support Curve25519, which is needed in Signal. Consequently, implementations like Signal Messenger end up compiling cryptographic primitives from C to JavaScript. Even if the desired primitives were available in the underlying platform, accessing them in a hostile environment is unsafe, since an attacker may have redefined them. Consequently, we developed our own libraries for cryptography and message encoding.

The ProScript Cryptography Library (PSCL) is a trusted cryptographic library implementing a variety of modern cryptographic primitives such as X25519, AES-CCM and BLAKE2. All of its primitives are fully type-checked without this affecting speed: in the majority of our benchmarks, PSCL is as fast as or faster than popular JavaScript cryptographic libraries like SJCL and MSR JavaScript Crypto, which do not even benefit from defensive type checking (Figure 3).

More crucially, PSCL functions used in ProScript code are detected by the ProScript compiler as it produces the applied pi model of the implementation, giving it the ability to convert each call to a cryptographic primitive to a call to the corresponding symbolic function in ProVerif. For example, if the ProScript compiler sees a call to PSCL's X25519 implementation, it will automatically translate it to a standard Diffie-Hellman construction in ProVerif.

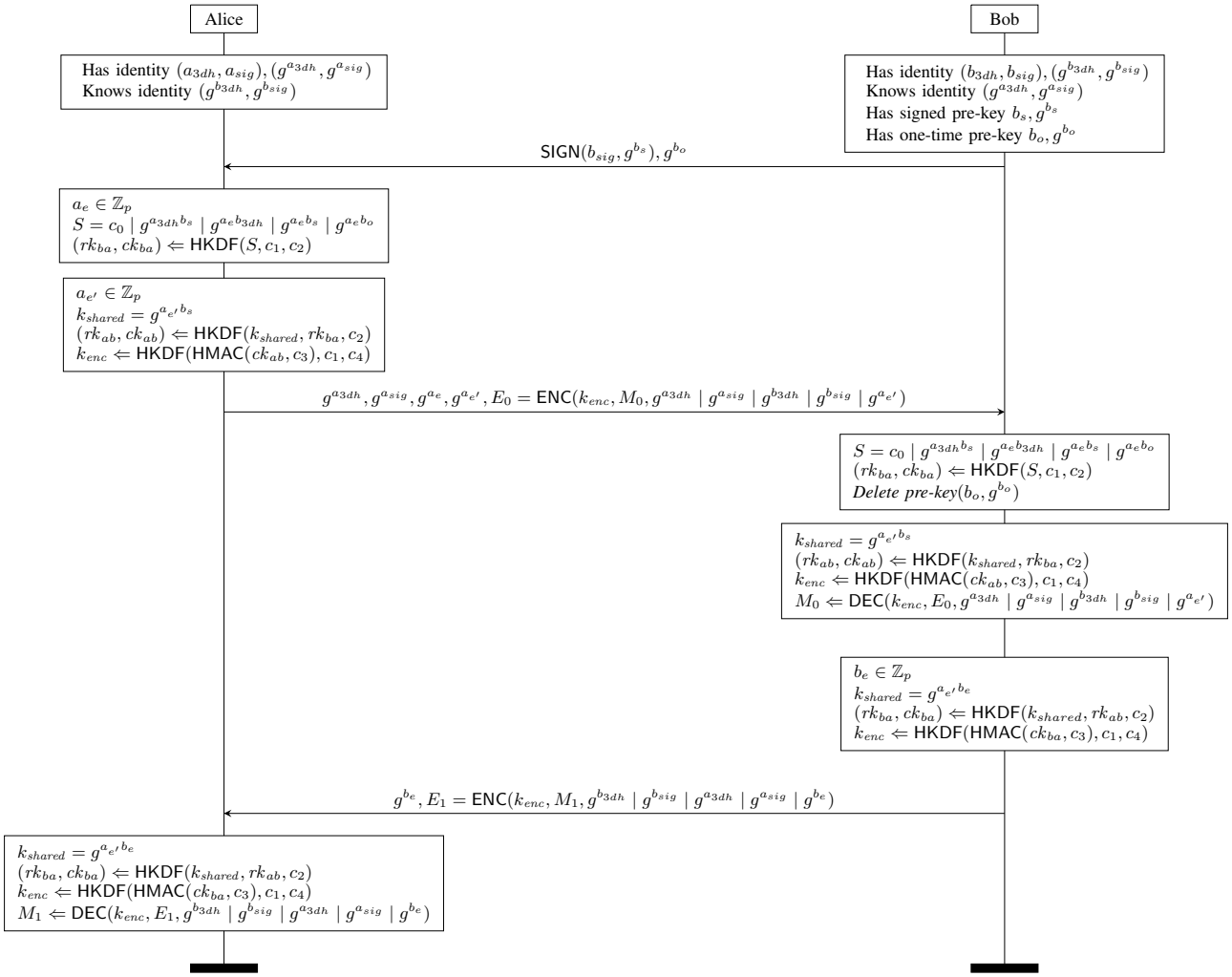

Figure 4: SP, a variant of Signal with minor differences. Alice requests a signed pre-key from Bob (via the server) and sends an initial message M_0. Bob accomplishes his side of the key exchange and obtains M_0. Bob later sends his reply M_1, illustrating the Axolotl ratchet post-AKE. We ignore the hash-based ratchet that occurs when two consecutive messages are sent in the same direction. c_i refers to various constants found throughout the protocol.

5. Implementing and Verifying SP

We describe SP, a variant of Signal Protocol that closely follows TextSecure version 3. We show how we implement and verify this protocol in our framework.

5.1. Protocol Overview

In SP, as illustrated in Figure 4, each client publishes a long-term Diffie-Hellman public key and a set of ephemeral Diffie-Hellman public keys (called *"pre-keys"*). These keys include both signed pre-keys, which can be reused for some period of time, and non-signed, one-time pre-keys, which are fresh at each session. To send a message to Bob, Alice retrieves Bob's long-term keys $(g^{b_{3dh}}, g^{b_{sig}})$, a signed pre-key g^{b_s} and a one-time pre-key g^{b_o}. She then chooses her own ephemeral g^{a_e}. A four-way Diffie-Hellman handshake is accomplished using Alice and Bob's long-term identity keys and their short-term ephemeral keys in order to derive the session secret S. The one-time pre-key is optional: when there remains no available one-time pre-key, the exchange is performed with a triple Diffie-Hellman handshake. An encryption key, k_{enc}, is then derived from S by Hash-Based Key Derivation (HKDF) [22] and the message M_0 is sent encrypted under the authenticated encryption scheme AES-GCM, with public and ephemeral keys as associated data: $\mathsf{ENC}(k, m, ad)$ means that m is encrypted with k and both the message m and the associated data ad are au-

thenticated. Subsequent messages in the conversation obtain authentication by chaining to S via a forward-ratcheting construction that also employs HKDF. Each sent message includes its own newly generated ephemeral public key and the protocol's double ratchet key refresh mechanism manages the key state by advancing key chaining with every message.

SP's forward and future secrecy goals are intended to make it so that the compromise of Alice or Bob's long-term keys allows for their impersonation but not for the decryption of their messages. The use of a signed initial ephemeral pre-key results in weaker forward secrecy guarantees for the first flight of messages from A to B: no forward secrecy is provided if both the long-term keys and pre-keys are leaked, although the guarantees for subsequent flights remain strong. If pre-keys are not signed, then the protocol only offers weak forward secrecy with respect to long-term key leakage. We note that the term "forward secrecy" can be confusing in a protocol like Signal, because a number of keys are at stake: long-term keys $((a_{3dh}, a_{sig}), (b_{3dh}, b_{sig}))$, signed pre-key b_s, one-time pre-key b_o, ephemeral keys $(a_e, a_{e'}, b_{e'})$, root keys $(rk_{ab}, ck_{ab}, rk_{ba}, ck_{ba})$ and message keys (k_{enc}). Any formal analysis of the protocol must precisely state which of these keys can be compromised and when.

Differences from other versions of Signal. An earlier version of the protocol, TextSecure Version 2, was cryptographically analyzed in previous work [2]. There are two key differences between SP and TextSecure Version 2.

Signed, Time-Based Pre-Keys. Version 2 uses a triple Diffie-Hellman handshake with one of a hundred pre-keys that Bob stores on the server (including a "last-resort" pre-key). TextSecure Version 3 and all subsequent versions of Signal, including SP, use a signed time-based pre-key, used in conjunction with an unsigned one-time pre-key in case it is available. Bob periodically replaces his signed pre-key (for example, once every week), which may be re-used until its replacement and refreshes his collection of unsigned one-time pre-keys. In SP, when one-time pre-keys are exhausted, no "last-resort" pre-key is used.

Stronger Identity Protection. Since Version 3, tag_n is expanded to include the long-term identities of the sender and recipient, which is not the case in Version 2. This provides a slightly stronger authentication guarantee in the rare case that the encryption keys for different pairs of users turns out to be the same.

In addition to these differences with Version 2, SP also differs from other variants of Signal in one key aspect. In SP, long-term identities are split into one Diffie-Hellman key pair and one signing key pair. In Signal, the same key pair is used for both operations, by applying an elliptic curve conversion from the Montgomery curve-based Diffie-Hellman key pair to obtain its twisted Edwards curve-based signing key pair equivalent. We choose to use separate keys instead, because in our cryptographic proof, we do not want to add a non-standard cryptographic assumption about the use of the same key in two independent cryptographic operations. In exchange for using standard cryptographic assumptions, we consider the cost of adding an extra 32 byte key to the protocol to be acceptable.

5.2. Protocol Implementation

To implement SP in ProScript, we must first deconstruct it into various protocol components: structures for managing keys and user states, messaging functions, APIs and top-level processes. ProScript is well-equipped to handle these protocol components in a way that lends itself to model extraction and verification. We break down our ProScript SP implementation into:

Types for State and Key Management. ProScript's type declaration syntax can be used to declare types for individual elements such as encryption keys but also for collections of elements such as a conversation party's state. These declarations allow for the construction of common data structures used in the protocol and also makes their management and modification easier in the extracted ProVerif models.

Messaging Interface. The ProScript implementation exposes the generic messaging API in a single global object. All interface access provides purely state-passing functionality.

Long-Term and Session States. Protocol functions take long-term and session states (S^a, T_n^{ab}) as input and return T_{n+1}^{ab}. S^a contains long-term values such identity keys, while T includes more session-dependent values such as `ephemerals`, containing the current ephemeral and chaining keys for the session context and `status`, indicating whether the application layer should perform a state update.

Internal Functions. Utility functionality, such as key derivation, can also be described as a series of pure functions that are not included in the globally accessible interface.

Top-Level Process. A top-level process can serve as a harness for testing the proper functioning of the protocol in the application layer. Afterwards, when this top-level process is described in the extracted ProVerif model, the implementer will be able to use it to define which events and security properties to query for.

Inferred Types in ProScript. ProScript type declarations allow for the easier maintenance of a type-checkable protocol implementation, while also allowing the ProScript compiler to translate declared types into the extracted ProVerif model. Defining a `key` as an array of 32 bytes will allow the ProScript compiler to detect all 32 byte arrays in the implementation as keys and type their usage accordingly.

5.3. Protocol Verification

We use ProVerif to verify the security goals of our extracted model by defining defining queries that accurately test the resilience of security properties against an active adversary. Under an active Dolev-Yao adversary, ProVerif was able to verify confidentiality, authenticity, forward secrecy and future secrecy for Alice and Bob initializing a session

Goals	Messages	Parties	Roles	Time
Secrecy	1	A, B	One	00h.04m.07s.
Secrecy	1	A, B	Two	00h.11m.17s.
Indist.	1	A, B	One	02h.06m.15s.
Authen.	1	A, B, M	One	00h.58m.19s.
Authen.	1	A, B, M	Two	29h.17m.39s.
Fo. Se.	1	A, B	One	00h.04m.14s.
KCI	1	A, B	One	00h.19m.20s.

Figure 5: Verification times for SP ProVerif models.

and exchanging two secret messages, with a compromised participant, Mallory, also being allowed to initialize sessions and exchange non-secret messages with Alice and Bob. Our analysis revealed two novel attacks: a key compromise impersonation attack and a replay attack, for which we propose a fix. Aside from these attacks, we were also able to model the previously documented Unknown Keyshare Attack [2].

Extracts of our compiled SP implementation are available online [23]. Models begin with type declarations followed by public constant declarations, equational relationships for cryptographic primitives, protocol functions, queries and relevant names and finally the top-level process with its associated queries.

The top-level process queries for security properties such as confidentiality, authenticity and forward secrecy between two roles: an initiator (e.g. Alice) who sends an initial message and thereby initializes an authenticated key exchange, and a responder (e.g. Bob) who receives the message and who may send a response. Some models include a third compromised identity, Mallory, who also communicates with Alice and Bob but while leaking her private keys to the attacker beforehand. In some instances, we also model parallel process executions where each identity (Alice, Bob and optionally Mallory) assumes both the role of the initiator and the responder. We informally call this latter scenario a "two-way role" model.

Secrecy and Indistinguishability. For every message considered in our protocol model, we define a secret constant M_n where M_1 is the initial message in a session. These secret values are then used as the plaintext for the encrypted messages sent by the principals. We show that an active attacker cannot retrieve a message's plaintext M_n using the query:

$$\mathsf{query}(\mathsf{attacker}(M_n)) \tag{1}$$

Similarly, we show indistinguishability using the query $\mathsf{query}(\mathsf{noninterf}(M_n))$.

Forward and Future Secrecy. We examine forward and future secrecy in Signal Protocol in multiple scenarios: the compromise of long-term keys and the compromise of message keys in two different types of message flights. In these scenarios, we need to model that keys are leaked after sending or receiving certain messages. We rely on ProVerif *phases* for that: intuitively, t represents a global clock, and processes occurring after the declaration of a phase t are active only during this phase.

We show that message M_1 remains secret by query (1) even if the long-term keys $(a_{3dh}, a_{sig}, b_{3dh}, b_{sig})$ are leaked after sending M_1. Furthermore, we can modify our ProVerif model to produce a sanity check: if responder Bob skips the signature check on g^{a_s}, ProVerif shows that an active attacker becomes capable of violating this forward secrecy property.

Next, we examine two different messaging patterns in the Double Ratchet algorithm and find that they approach forward and future secrecy differently:

- **Single-Flight Pattern.** In this scenario, Alice sends Bob a number of messages M_n and M_{n+1} where $n > 1$ and does not receive a response. In this scenario, Bob's lack of response does not allow Alice to obtain a fresh ephemeral key share g^{b_e} required to establish a new k_{shared} in T^{ab}_{n+1} to be used for M_{n+1}, so Alice just updates the key ck_{ab} by hashing it. If Alice's session state T^{ab}_{n+1}, (which, recall, contains ${a_e}^{n+1}$ and (rk_{ab}, ck_{ab}) for M_{n+1}), is leaked, then M_n remains secret (forward secrecy). Obviously, to take advantage of this property in case of compromise, the keys (rk_{ab}, ck_{ab}) for M_n must have been appropriately deleted, which is delicate when messages are received out-of-order: if $M_{n_1}, \ldots, M_{n_k}$ $(n_1 < \ldots < n_k)$ have been received, the receiver should keep the chaining key ck_{ab} for M_{n_k+1} and the encryption keys k_{enc} for the messages M_i not received yet with $i < n_k$. If T^{ab}_n is leaked, then M_{n+1} is not secret, so no future secrecy is obtained.
- **Message-Response Pattern.** In this scenario, Alice sends Bob a single message M_n where $n > 1$ and receives a response M_{n+1} before sending M_{n+2}. Upon receiving M_{n+1}, Alice will be able to derive a fresh $k_{shared} = g^{{a_e}^{n+2}{b_e}^{n+1}}$. As a result, if T^{ab}_{n+2} is leaked, then M_n remains secret (forward secrecy) and if T^{ab}_n is leaked after M_{n+1} is received, then M_{n+2} remains secret (future secrecy).

Message Authenticity. Signal Protocol relies on a Trust-on-First-Use (TOFU) authentication model: Alice assumes that Bob's advertised identity key is authenticated and untampered with and employs it as such until an event causes the trust of the key to be put in question, such as a sudden identity key change or an out of band verification failure. We model TOFU by embedding Alice and Bob's identity keys into each other's initial states. We are then free to model for message authenticity: informally, if B receives a message M from A, we want A to have sent M to B. In ProVerif, we can specify two events: $\mathsf{Send}(A, B, M)$, which means that A sends M to B and $\mathsf{Recv}(A, B, M)$, which means that B receives M from A. We can then formalize the correspondence

$$\mathsf{event}(\mathsf{Recv}(A, B, M)) \Longrightarrow \mathsf{event}(\mathsf{Send}(A, B, M)) \tag{2}$$

which checks if for all $\mathsf{Recv}(A, B, M)$ events, it must be the case that a $\mathsf{Send}(A, B, M)$ event has also been executed.

ProVerif succeeds in proving correspondence (2) using public keys A and B. While this implies the desired property when the relation between the public keys and the identity of the principals is bijective, a limitation of this approach is that the identities of the principals are only expressed in terms of keys and not as a more personally-linked element, such as for example a phone number. Therefore, we cannot formally express stronger identity binding as part of the protocol model. This point leads to the Unknown Key Share Attack first reported for Signal Protocol Version 2 [2]: if an adversary can register the public keys $(g^{b_{3dh}}, g^{b_{sig}})$ of B as public keys of C and A sends a message to C, then C can forward this message to B and B will accept it as coming from A, since B and C have the same public keys.

No Replays. This property is similar to message authenticity, but uses an injective correspondence instead, which means that each execution of $\mathsf{Recv}(A, B, M)$ corresponds to a distinct execution of $\mathsf{Send}(A, B, M)$:

$$\mathsf{inj\text{-}event}(\mathsf{Recv}(A, B, M)) \Longrightarrow \mathsf{inj\text{-}event}(\mathsf{Send}(A, B, M))$$

When a optional one-time pre-key is involved in the initial session handshake, ProVerif shows that the injective correspondence holds for the first message in the conversation. However, when this optional one-time pre-key is not used, a replay attack is detected. Signal Protocol Version 3 will accept a Diffie-Hellman handshake that only employs identity keys and signed pre-keys, both of which are allowed to be reused across sessions. This reuse is what makes a replay attack possible. We propose a fix for this issue by having clients keep a cache of the ephemeral keys used by the sender of received messages, associated with that sender's identity key. We are able to expand our event queries in ProVerif to account for this fix by showing the non-injective correspondence of the Send and Recv events with added ephemeral keys. Coupled with a caching of ephemeral keys, we can ensure that the Recv event is only executed once per ephemeral key. Hence, the injective correspondence is implied by the non-injective correspondence.

Key Compromise Impersonation (KCI). We present a novel key compromise impersonation attack: to detect KCI, we consider a scenario in which Alice or Bob's keys are compromised and test again for authenticity of messages received by the compromised principal. When Alice or Bob's long-term secret key is compromised, ProVerif shows that message authenticity still holds. However, when Bob's signed pre-key is also compromised, ProVerif finds an attack against message authenticity. This is a novel key compromise impersonation attack: when the adversary has Bob's signed pre-key s, he can choose x and x' and compute the session keys using Alice's and Bob' public keys $(g^{a_{3dh}}, g^{a_{sig}})$ and $(g^{b_{3dh}}, g^{b_{sig}})$ and Bob's one time pre-key g^o and send his own message in Alice's name. This message is accepted by Bob as if it came from Alice: the event $\mathsf{Recv}(A, B, M)$ is executed without having executed $\mathsf{Send}(A, B, M)$.

5.4. Other Protocols: OTR

To obtain higher confidence in ProScript's usability as an implementation and verification framework, we also implemented Off-the-Record (OTR) Messaging Version 2, which has been shown to include various vulnerabilities discovered under finite-state analysis by Bonneau and Morrison [24]. OTR [17] uses a SIGMA-based key exchange [25] and a ratcheting scheme in order to achieve confidentiality, authentication, perfect forward secrecy and deniability. During model verification, known vulnerabilities in OTR Version 2 were automatically detected:

Version Rollback Attack. Since the communication of which OTR versions are supported by a client was performed before authentication occurred, it is possible to maliciously force users to adopt an outdated version of OTR.

Unknown Key-Share Attack. Since OTR's authenticated key exchange lacks sufficient information identifying Alice's intended recipient, Mallory can forward handshake messages in order to successfully trick Alice into establishing a session with Bob, whereas Alice intended to establish a session with Mallory.

Message Integrity Attack. If Alice receives a message from Bob where Bob uses new ratchet keys, she will publish her MAC keys for her previous message. This is an intentional feature meant to provide deniability once both parties are certain that a certain step of ratchet keys are obsolete. However, Mallory can simply block this outgoing message, use it to learn the MAC key for Alice's previous message and commit a forgery under those keys. From Bob's perspective, it would appear that he has received a delayed but nevertheless valid message from Alice.

Integrating Symbolic Verification into the Development Cycle. Human-readability of the automatically compiled ProVerif model is key to our verification methodology. In the case of a query failure, users can opt to modify their implementation and recompile into a new model, or they can immediately modify the model itself and re-test for security queries within reasonable model verification times. For example, if an implementer wants to test the robustness of a passing forward secrecy query, they can disable the signature verification of signed pre-keys by changing a single line in the model, causing the client to accept any pre-key signature.

5.5. Results of Key Compromise

- $((a_{3dh}, a_{sig}), (b_{3dh}, b_{sig}))$ Alice and Bob lose authenticity guarantees for all subsequent messages. Keys in `secrets` are vulnerable to a device-level attacker due to their permanent storage and reuse across all sessions.
- $(a_e, g^{a_e}, b_o, g^{b_o})$ In Signal Protocol Version 2, and only in the case of an earlier long-term key compromise (a or b), Alice and Bob lose confidentiality and integrity guarantees for the subsequent message.
- $ck_{ab} \in T_n^{ab}$ Immediate loss of confidentiality and integrity for the subsequent message.

6. Cryptographic Proofs with CryptoVerif

To complement the results obtained in the symbolic model using ProVerif, we use the tool CryptoVerif [5] in order to obtain security proofs in the computational model. This model is much more realistic: messages are bitstrings; cryptographic primitives are functions from bitstrings to bitstrings; the adversary is a probabilistic Turing machine. CryptoVerif generates proofs by sequences of games [26], [27], like those written manually by cryptographers, automatically or with guidance of the user.

The computational model is more realistic, but it also makes it more difficult to mechanize proofs. For this reason, CryptoVerif is less flexible and more difficult to use than ProVerif, and our results in the computational model are more limited. We model only one message of the protocol (in addition to the pre-keys), so we do not prove properties of the ratcheting algorithm. Considering several data messages exceeds the current capabilities of CryptoVerif—the games become too big.

Rather than directly using models generated from our ProScript code, we manually rewrite the input scripts of CryptoVerif, for two main reasons:

- The syntax of the protocol language of CryptoVerif differs slightly from that of ProVerif. We plan to overcome this difficulty in the future by modifying the syntax of CryptoVerif so that it is compatible with ProVerif.
- The kinds of models that are easy to verify using CryptoVerif differ from those that are easy for ProVerif; therefore, even if the source syntax were the same, we would still need to adapt our compiler to generate specialized models that would be more conducive to CryptoVerif's game-based proofs.

6.1. Assumptions

We make the following assumptions on the cryptographic primitives:

- The elliptic curve Ec25519 satisfies the gap Diffie-Hellman (GDH) assumption [28]. This assumption means that given g, g^a, and g^b for random a, b, the adversary has a negligible probability to compute g^{ab} (computational Diffie-Hellman assumption), even when the adversary has access to a decisional Diffie-Hellman oracle, which tells him given G, X, Y, Z whether there exist x, y such that $X = G^x$, $Y = G^y$, and $Z = G^{xy}$. When we consider sessions between a participant and himself, we need the square gap Diffie-Hellman variant, which additionally says that given g and g^a for random a, the adversary has a negligible probability to compute g^{a^2}. This assumption is equivalent to the GDH assumption when the group has prime order [29], which is true for Ec25519 [30]. We also added that $x^y = x'^y$ implies $x = x'$ and that $x^y = x^{y'}$ implies $y = y'$, which hold when the considered Diffie-Hellman group is of prime order.
- Ed25519 signatures, used for signing pre-keys, are unforgeable under chosen-message attacks (UF-CMA) [31].
- The functions

$$x_1, x_2, x_3, x_4 \mapsto \mathsf{HKDF}(x_1 \| x_2 \| x_3 \| x_4, c_1, c_2)$$
$$x_1, x_2, x_3 \mapsto \mathsf{HKDF}(x_1 \| x_2 \| x_3, c_1, c_2)$$
$$x, y \mapsto \mathsf{HKDF}(x, y, c_2)$$
$$x \mapsto \mathsf{HKDF}(x, c_1, c_4)$$

  are independent random oracles, where x, y, x_1, x_2, x_3, x_4, and c_1 are 256-bit long. We further justify this assumption in the Appendix: there, we show that these functions are indifferentiable [32] from independent random oracles, assuming that the compression function underlying SHA256 is a random oracle. (The considered HKDF function [22] is defined from HMAC-SHA256, which is itself defined from SHA256.)
- HMAC-SHA256 is a pseudo-random function (PRF) [33]. This assumption is used for $\mathsf{HMAC}(ck_{ab}, \cdot)$ and $\mathsf{HMAC}(ck_{ba}, \cdot)$.
- The encryption scheme ENC, which is AES-GCM, is a secure authenticated encryption with associated data (AEAD). More precisely, it is indistinguishable under chosen plaintext attacks (IND-CPA) and satisfies ciphertext integrity (INT-CTXT) [34], [35].

CryptoVerif provides a library that predefines the most common cryptographic assumptions, so that the user does not have to write them for each protocol. In our work, we had to adapt these predefined assumptions to our specific needs: the GDH assumption is predefined, but the square GDH variant is not; unary random oracles are predefined, but we also needed binary, ternary, and 4-ary ones; predefined PRFs, SUF-CMA MACs, and IND-CPA encryption schemes use a key generation function, while in our schemes the key is a plain random bitstring, without a key generation function. Adapting the definition of primitives did not present any major difficulty. As mentioned in § 5.1, we had to make one modification to the original Signal Protocol, for it to be provable in the computational model: we use different keys for the elliptic curve Diffie-Hellman and elliptic curve signatures. It is well-known that using the same keys for several cryptographic primitives is undesirable, as proving security requires a joint security assumption on the two primitives in this case. Therefore, we assume each protocol participant to have two key pairs, one for Diffie-Hellman and one for signatures. This problem remains undetected in a symbolic analysis of the protocol.

6.2. Protocol Model

We model SP as a process in the input language of CryptoVerif, which is similar to the one of ProVerif. We consider simultaneously the protocol of Figure 4 and the version without the optional one-time pre-key b_o. As mentioned above, we consider only one message in each session. Our threat model includes an untrusted network, malicious principals, and long-term key compromise, as mentioned in § 2. It does not include session state compromise, which is less useful with a single message.

At a high level, we use the same messaging API as in § 2. However, to make verification easier for CryptoVerif, we specify a lower-level interface. We consider two honest principals Alice and Bob, and define separate processes for Alice interacting with Bob, with herself, or with a malicious participant, Bob interacting with Alice, and Bob interacting with himself or a malicious participant, as well as similar processes with the roles of Alice and Bob reversed. The adversary can then implement the high-level interface of § 2 from this lower-level interface: the adversary is supposed to implement the malicious principals (including defining keys for them) and to call the low-level interface processes to run sessions that involve the honest principals Alice and Bob.

We make two separate proofs: In the first one, we prove the security properties for sessions in which Bob generates pre-keys and runs the protocol with Alice. (Other protocol sessions exist in parallel as described above; we do not prove security properties for them. For sessions for which we do not prove security properties, we give to the adversary the ephemeral a'_e and the key rk_{ba} or rk_{ba} and let the adversary encrypt and MAC the message himself, to reduce the size of our processes.) In the second one, we prove the security properties for sessions in which Alice generates pre-keys and runs the protocol with herself. Bob is included in the adversary in this proof. The security for sessions in which Alice generates pre-keys and runs the protocol with Bob follows from the first proof by symmetry. The security for sessions in which Bob generates pre-keys and runs the protocol with himself follows from the second proof. The other sessions do not satisfy security properties since they involve the adversary. (They must still be modeled, as they could break the protocol if it were badly designed.) Therefore, these two proofs provide all desired security properties.

6.3. Security Goals

We consider the following security goals from § 2:

Message Authenticity, No Replays, and Key Compromise Impersonation (KCI). These properties are modeled by correspondences as in ProVerif (§ 5.3). For key compromise impersonation, we consider the compromise of the long-term Diffie-Hellman and signature keys of Bob, and prove again message authenticity. We do not consider the compromise of the signed pre-key since we already know from the symbolic analysis that there is an attack in this case.

Computational Indistinguishability. If A randomly chooses between two messages M_0, M_1 of the same length and sends one of them to B, then the adversary has a negligible probability of guessing which of the two messages was sent. In our model, this is formalized by choosing a random bit $secb \in \{0, 1\}$; then A sends message M_b to B, and we show that the bit $secb$ remains secret, with the query $\mathbf{secret}\ secb$.

Forward Secrecy. This is proved exactly like indistinguishability, but with an additional oracle that allows the adversary to obtain the secret keys of the principals, thus compromising them.

Goals	Parties	Running Time
Forward Secrecy	A, B, M	3 min. 58 sec.
Forward Secrecy	A, M	7 min. 04 sec.
KCI	A, B, M	3 min. 15 sec.
Others	A, B, M	4 min. 15 sec.
Others	A, M	3 min. 35 sec.

Figure 6: Verification times for SP CryptoVerif models, without anti-replay countermeasure. The runtimes with the anti-replay countermeasure are of the same order of magnitude. Tested using CryptoVerif 1.24.

We do not consider future secrecy since we have a single message. We do not consider secrecy since we directly deal with the stronger property of indistinguishability.

6.4. Results

CryptoVerif proves message authenticity, absence of key compromise impersonation attacks (when the long-term keys of Bob are compromised), indistinguishability, and forward secrecy, but cannot prove absence of replays. This is due to the replay attack mentioned in § 5.3. Since this attack appears only when the optional one-time pre-key is omitted, we separate our property into two: we use events $\mathsf{Send}(A, B, M)$ and $\mathsf{Recv}(A, B, M)$ for the protocol with optional pre-key and events $\mathsf{Send3}(A, B, M)$ and $\mathsf{Recv3}(A, B, M)$ for the protocol without optional pre-key. CryptoVerif then proves

$$\mathsf{inj\text{-}event}(\mathsf{Recv}(A, B, M)) \Longrightarrow \mathsf{inj\text{-}event}(\mathsf{Send}(A, B, M))$$
$$\mathsf{event}(\mathsf{Recv3}(A, B, M)) \Longrightarrow \mathsf{event}(\mathsf{Send3}(A, B, M))$$

which proves message authenticity and no replays when the one-time pre-key is present and only message authenticity when it is absent. This is the strongest we can hope for the protocol without anti-replay countermeasure.

With our anti-replay countermeasure (§5.3), CryptoVerif can prove the absence of replays, thanks to a recent extension that allows CryptoVerif to take into account the replay cache in the proof of injective correspondences, implemented in CryptoVerif version 1.24. Our CryptoVerif proofs have been obtained with some manual guidance: we indicated the main security assumptions to apply, instructed CryptoVerif to simplify the games or to replace some variables with their values, to make terms such as $m^a = m^b$ appear. The proofs were similar for all properties.

7. A Verified Protocol Core for Cryptocat

We now describe how we can rewrite Cryptocat to incorporate our ProScript implementation of SP. We deconstruct the Cryptocat JavaScript code into the following components, as advocated in Figure 2.

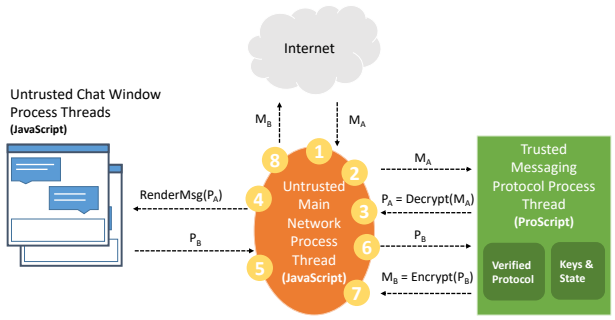

Figure 7: Cryptocat Architecture: isolating verified and untrusted components in Electron apps within separate processes.

1) **Unverified JavaScript Application** This component, which comprises the majority of the code, manages the user's state, settings, notifications, graphical interface and so on. It is connected to the protocol only via the ability to call exposed protocol functions (as documented in § 2.1). We adopt certain assumptions regarding the unverified JavaScript application, for example that it will not modify the protocol state outside of passing it through the protocol implementation interface.
2) **Verified Protocol Implementation** This component is written in ProScript and resides in a separate namespace, functioning in a purely state-passing fashion. Namely, it does not store any internal state or make direct network calls. This implementation is type-checked, and automatically verified every time it is modified.
3) **Trusted Library** This component provides cryptographic functionality. A goal is to include modern cryptographic primitives (X25519, AES-CCM) and provide type-checking assurances without affecting speed or performance.

This layered architecture is essential for our verification methodology, but is quite different from other messaging applications. For example, the Signal Desktop application is a Chrome browser application also written in JavaScript [36]. Parts of the protocol library are compiled from C using Emscripten, presumably for performance, parts are taken from third-party libraries, and other protocol-specific code is written in JavaScript. The resulting code (1.5MB, 39Kloc) is quite hard to separate into components, let alone verify for security. We hope that our layered approach can lead to verified security guarantees without sacrificing performance or maintainability.

7.1. Isolating Verified Code

We build Cryptocat using Electron [37], a framework for JavaScript desktop applications. Electron is built on top of the Node.js JavaScript runtime and the Chromium web renderer. Electron has recently gained strong adoption: popular applications such as Visual Studio Code, Atom, Slack, WordPress and WhatsApp Desktop are all built on top of Electron.

By default, Electron allows applications to load any Node.js low-level module, which can in turn perform dangerous operations like accessing the file system and exfiltrate data over the network. Since all Node.js modules in a single process run within the same JavaScript environment, malicious or buggy modules can tamper with other modules via prototype poisoning or other known JavaScript attack vectors. Consequently, all Electron apps, including other desktop Signal implementations like WhatsApp and Signal messenger effectively include all of Electron and Node.js into their trusted computing base (TCB).

We propose a two-pronged approach to reduce this TCB.

Language-Based Isolation. Since ProScript is a subset of Defensive JavaScript, ProScript protocol code is isolated at the language level from other JavaScript code running within the same process, even if this code uses dangerous JavaScript features such as prototype access and modification. To ensure this isolation, ProScript code must not call any external (untyped) libraries.

Process Thread Isolation. We exploit features of Electron in order to isolate components of our application in different CPU threads as seen in Figure 7. When a message arrives on the network (1), the main network thread can only communicate with our Protocol TCB using a restrictive inter-process communication API. The TCB then uses its internal verified protocol functionality and state management to return a decryption of the message (3), which is then forwarded again via IPC to the chat window (4), a third separate CPU thread which handles message rendering. Furthermore, the TCB process is disallowed from loading any Node.js modules.

In particular, the network process is isolated from the chat media rendering process; neither ever obtain access to the key state or protocol functionality, which are all isolated in the ProScript protocol process. When Bob responds, a similar IPC chain of calls occurs in order to send his reply back to Alice (5, 6, 7, 8). Even if an error in the rendering code or in the XML parser escalated into a remote takeover of the entire web renderer, the calls to the protocol TCB would be restricted to those exposed by the IPC API. However, these isolation techniques only protect the ProScript code within our application when executed within a correct runtime framework. None of these techniques can guard against bugs in V8, Node.js, or Electron, or against malicious or buggy Node.js or Electron modules loaded by the application.

7.2. Performance and Limitations

Although we have verified the core protocol code in Cryptocat and tried to isolate this code from unverified code, the following limitations still apply: we have not formally verified the soundness of the cryptographic primitives themselves, although writing them in Defensive JavaScript does provide type safety. We have also not formally verified the Electron framework's isolation code. Similarly, we do not claim any formal verification results on the V8 JavaScript runtime or on the Node.js runtime. Therefore, we rely on

a number of basic assumptions regarding the soundness of these underlying components. Cryptocat's successful deployment provides a general guideline for building, formally verifying, and isolating cryptographic protocol logic from the rest of the desktop runtime. Designing better methods for implementing and isolating security-critical components within Electron apps with a minimal TCB remains an open problem. Despite the strong programming constraints imposed by our verification architecture, we find that Cryptocat is able to perform similarly to mainstream desktop messaging applications that do not offer end-to-end encryption, such as Skype. In order to benefit from automatic model translation of our ProScript protocol implementation (as described in §4.4), we use PSCL as our cryptography library, which allows us to easily handle even large file encryptions (200+MB) for Cryptocat's file sharing feature. Our application is available for Windows, Linux and Mac, and currently serves over 20,000 users weekly. It is capable of handling multi-device provisioning for users with soft device revocation and device authentication. We also support video messaging, file sharing and similar usability features, which all exploit the isolation and formal verification methods described in this paper.

8. Related Work

Extracting Protocol Models from Running Code. There have been previous attempts [38] to extract ProVerif models from typed JavaScript, such as *DJS2PV* [7]. However, DJS2PV was only tested on small code examples: attempting to translate a complete implementation such as Signal Protocol resulted in a 3,800 line model that attempts to precisely account for the heap, but could not verify due to an exploding state space. Previous efforts such as FS2PV [20] avoided this problem by choosing a purely functional source language that translated to simpler pi calculus scripts. We adopt their approach in ProScript to generate briefer, more readable models.

Type Systems for JavaScript. TypeScript [39], Flow [40], Defensive JavaScript and TS* [41] all define type systems that can improve the security of JavaScript programs. The type system in ProScript primarily serves to isolate protocol code from untrusted application and to identify a subset of JavaScript that can be translated to verifiable models.

Formal Analysis of Web Security Protocols. Tools like WebSpi [42] and AuthScan [43] have been used to verify the security of web security protocols such as OAuth. An expressive web security model has also been used to build manual proofs for cryptographic web protocols such as BrowserID [44]. These works are orthogonal to ProScript and their ideas can potentially be used to improve our target ProVerif models.

Analysis of Secure Messaging Protocols. Unger et al. survey previous work on secure messaging [10]. We discuss three recent closely-related works here.

Future secrecy was formalized by Cohn-Gordon et al. as "post-compromise security" [45]. Our symbolic formulation is slightly different since it relies on the definition of protocol phases in ProVerif.

Cryptographic security theorems and potential unknown key-share attacks on TextSecure Version 2 were presented by Frosch et al. [2]. In comparison to that work, our analysis covers a variant of TextSecure Version 3, our analysis is fully mechanized, and we address implementation details. Our CryptoVerif model only covers a single message, but we consider the whole protocol at once, while they prove pieces of the protocol separately. Like we do, they consider that HKDF is a random oracle. We further justify this assumption by an indifferentiability proof.

More recently and in parallel with this work, Cohn-Gordon et al. [46] prove, by hand, that the message encryption keys of Signal are secret in the computational model, in a rich compromise scenario, under assumptions similar to ours. Thereby, they provide a detailed proof of the properties of the double ratcheting mechanism. However, they do not model the signatures of the signed pre-keys, and they do not consider key compromise impersonation attacks or replay attacks or other implementation-level details. In contrast to their work, our computational proof is mechanized, but limited to only one message.

9. Conclusion and Future Work

Drawing from existing design trends in modern cryptographic web application, we have presented a framework that supports the incremental development of custom cryptographic protocols hand-in-hand with formal security analysis. By leveraging state-of-the-art protocol verification tools and building new tools, we showed how many routine tasks can be automated, allowing the protocol designer to focus on the important task of analyzing her protocol for sophisticated security goals against powerful adversaries.

We plan to continue to develop and refine ProScript by evaluating how it is used by protocol designers, in the spirit of an open source project. All the code and models presented in this paper, and a full version of this paper are available online [23]. Proving the soundness of translation from ProScript to ProVerif, by relating the source JavaScript semantics to the applied pi calculus, remains future work. The process of transforming the compiled model to a verified CryptoVerif script remains a manual task, but we hope to automate this step further, based on new and upcoming developments in CryptoVerif.

Finally, a word of caution: a protocol written in ProScript and verified with ProVerif or CryptoVerif does not immediately benefit from assurance against all possible attacks. Programming in ProScript imposes a strict discipline by requiring defensive self-contained code that is statically typed and can be translated to a verifiable model and subsequent verification can be used to eliminate certain well-defined classes of attacks. We believe these checks can add confidence to the correctness of a web application, but they do not imply the absence of security bugs, since we still have a large trusted computing base. Consequently, improving the robustness and security guarantees of runtime frameworks

such as Electron, Node.js, and Chromium, remains an important area of future research.

Acknowledgments. This work was funded by the following grants: ERC CIRCUS, EU NEXTLEAP, and ANR AJACS.

References

[1] K. Bhargavan, A. Lavaud, C. Fournet, A. Pironti, and P. Strub, "Triple handshakes and cookie cutters: Breaking and fixing authentication over TLS," in *IEEE Symposium on Security & Privacy (Oakland)*, 2014, pp. 98–113.

[2] T. Frosch, C. Mainka, C. Bader, F. Bergsma, J. Schwenk, and T. Holz, "How secure is TextSecure?" in *IEEE European Symposium on Security and Privacy (Euro S&P)*, 2016.

[3] B. Blanchet, "Modeling and verifying security protocols with the applied pi calculus and ProVerif," *Foundations and Trends in Privacy and Security*, vol. 1, no. 1–2, pp. 1–135, Oct. 2016.

[4] D. Dolev and A. C. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–207, 1983.

[5] B. Blanchet, "A computationally sound mechanized prover for security protocols," *IEEE Transactions on Dependable and Secure Computing*, vol. 5, no. 4, pp. 193–207, 2008.

[6] K. Bhargavan, A. Delignat-Lavaud, and S. Maffeis, "Language-based defenses against untrusted browser origins," in *USENIX Security Symposium*, 2013, pp. 653–670.

[7] K. Bhargavan, A. Delignat-Lavaud, and S. Maffeis, "Defensive JavaScript - building and verifying secure web components," in *Foundations of Security Analysis and Design (FOSAD VII)*, 2013, pp. 88–123.

[8] M. Abadi and C. Fournet, "Mobile values, new names, and secure communication," in *28th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'01)*. London, United Kingdom: ACM Press, Jan. 2001, pp. 104–115.

[9] H. Krawczyk, "HMQV: A High-performance Secure Diffie-Hellman Protocol," in *International Conference on Advances in Cryptology (CRYPTO)*, ser. Lecture Notes in Computer Science, V. Shoup, Ed., vol. 3621. Springer, 2005, pp. 546–566.

[10] N. Unger, S. Dechand, J. Bonneau, S. Fahl, H. Perl, I. Goldberg, and M. Smith, "SoK: Secure Messaging," in *IEEE Symposium on Security & Privacy (Oakland)*, 2015.

[11] N. Durov, "Telegram MTProto protocol," 2015, https://core.telegram.org/mtproto.

[12] O. Schirokauer, "The number field sieve for integers of low weight," *Mathematics of Computation*, vol. 79, no. 269, pp. 583–602, 2010.

[13] D. Gillmor, "Negotiated Finite Field Diffie-Hellman Ephemeral Parameters for Transport Layer Security (TLS)," 2016, IETF RFC 7919.

[14] K. Bhargavan, A. Delignat-Lavaud, and A. Pironti, "Verified contributive channel bindings for compound authentication," in *Proceedings of the ISOC Network and Distributed System Security Symposium (NDSS '15)*, Feb 2015. [Online]. Available: http://antoine.delignat-lavaud.fr/doc/ndss15.pdf

[15] A. Rad and J. Rizzo, "A 2ˆ64 attack on Telegram, and why a super villain doesn't need it to read your telegram chats." 2015.

[16] J. Jakobsen and C. Orlandi, "On the cca (in)security of mtproto," Cryptology ePrint Archive, Report 2015/1177, 2015, http://eprint.iacr.org/2015/1177.

[17] N. Borisov, I. Goldberg, and E. A. Brewer, "Off-the-record communication, or, why not to use PGP," in *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society, WPES 2004, Washington, DC, USA, October 28, 2004*, V. Atluri, P. F. Syverson, and S. D. C. di Vimercati, Eds. ACM, 2004, pp. 77–84. [Online]. Available: http://doi.acm.org/10.1145/1029179.1029200

[18] N. Wilcox, Z. Wilcox-O'Hearn, D. Hopwood, and D. Bacon, "Report of Security Audit of Cryptocat," 2014, https://leastauthority.com/blog/least_authority_performs_security_audit_for_cryptocat.html.

[19] P. A. Gardner, S. Maffeis, and G. D. Smith, "Towards a program logic for JavaScript," *SIGPLAN Not.*, vol. 47, no. 1, pp. 31–44, Jan. 2012. [Online]. Available: http://doi.acm.org/10.1145/2103621.2103663

[20] K. Bhargavan, C. Fournet, A. D. Gordon, and S. Tse, "Verified interoperable implementations of security protocols," *ACM Transactions on Programming Languages and Systems*, vol. 31, no. 1, 2008.

[21] Joyent Inc. and the Linux Foundation, "Node.js," 2016, https://nodejs.org/en/.

[22] H. Krawczyk, "Cryptographic extraction and key derivation: The HKDF scheme," in *Advances in Cryptology (CRYPTO)*, ser. Lecture Notes in Computer Science. Springer, 2010, vol. 6223, pp. 631–648.

[23] N. Kobeissi, "SP code repository," https://github.com/inria-prosecco/proscript-messaging, February 2017.

[24] J. Bonneau and A. Morrison, "Finite State Security Analysis of OTR Version 2," 2006.

[25] H. Krawczyk, "SIGMA: The 'SIGn-and-MAc' approach to authenticated Diffie-Hellman and its use in the IKE-protocols," in *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, ser. Lecture Notes in Computer Science, vol. 2729. Springer, 2003, pp. 400–425.

[26] V. Shoup, "Sequences of games: a tool for taming complexity in security proofs," IACR Cryptology ePrint Archive, 2004, http://eprint.iacr.org/2004/332.

[27] M. Bellare and P. Rogaway, "The security of triple encryption and a framework for code-based game-playing proofs," in *Advances in Cryptology (Eurocrypt)*, ser. Lecture Notes in Computer Science, S. Vaudenay, Ed., vol. 4004. Springer, May 2006, pp. 409–426.

[28] T. Okamoto and D. Pointcheval, "The gap-problems: a new class of problems for the security of cryptographic schemes," in *Practice and Theory in Public Key Cryptography (PKC)*, ser. Lecture Notes in Computer Science, K. Kim, Ed., vol. 1992. Springer, 2001, pp. 104–118.

[29] A. Fujioka and K. Suzuki, "Designing efficient authenticated key exchange resilient to leakage of ephemeral secret keys," in *Topics in Cryptology (CT-RSA)*, ser. Lecture Notes in Computer Science, A. Kiayias, Ed., vol. 6558. Springer, 2011, pp. 121–141.

[30] D. J. Bernstein, "Curve25519: New Diffie-Hellman speed records," in *Public Key Cryptography (PKC)*, 2006, pp. 207–228.

[31] S. Goldwasser, S. Micali, and R. Rivest, "A digital signature scheme secure against adaptive chosen-message attacks," *SIAM Journal of Computing*, vol. 17, no. 2, pp. 281–308, April 1988.

[32] J.-S. Coron, Y. Dodis, C. Malinaud, and P. Puniya, "Merkle-Damgård revisited: How to construct a hash function," in *Advances in Cryptology (CRYPTO)*, ser. Lecture Notes in Computer Science, vol. 3621. Springer, 2005, pp. 430–448.

[33] M. Bellare, "New proofs for NMAC and HMAC: Security without collision-resistance," in *Advances in Cryptology (CRYPTO)*, ser. Lecture Notes in Computer Science, C. Dwork, Ed., vol. 4117. Springer, 2006, pp. 602–619.

[34] D. A. McGrew and J. Viega, "The security and performance of the Galois/Counter Mode (GCM) of operation," in *Progress in Cryptology - INDOCRYPT 2004*, ser. Lecture Notes in Computer Science, A. Canteaut and K. Viswanathan, Eds., vol. 3348. Chennai, India: Springer, Dec. 2004, pp. 343–355.

[35] P. Rogaway, "Authenticated-encryption with associated-data," in *Ninth ACM Conference on Computer and Communications Security (CCS-9)*. Washington, DC: ACM Press, Nov. 2002, pp. 98–107.

[36] Open Whisper Systems, "Signal for the browser," 2015, https://github.com/WhisperSystems/Signal-Browser.

[37] GitHub, "Electron framework," 2016, http://electron.atom.io/.

[38] M. Avalle, A. Pironti, R. Sisto, and D. Pozza, "The Java SPI framework for security protocol implementation," in *Availability, Reliability and Security (ARES), 2011 Sixth International Conference on*, Aug 2011, pp. 746–751.

[39] G. Bierman, M. Abadi, and M. Torgersen, "Understanding TypeScript," in *ECOOP 2014 Object-Oriented Programming*, ser. Lecture Notes in Computer Science, R. Jones, Ed., vol. 8586. Springer, 2014, pp. 257–281.

[40] Facebook Inc., "Flow, a static type checker for JavaScript," http://flowtype.org/docs/about-flow.html.

[41] C. Fournet, N. Swamy, J. Chen, P.-E. Dagand, P.-Y. Strub, and B. Livshits, "Fully abstract compilation to JavaScript," *SIGPLAN Not.*, vol. 48, no. 1, pp. 371–384, Jan. 2013. [Online]. Available: http://doi.acm.org/10.1145/2480359.2429114

[42] C. Bansal, K. Bhargavan, A. Delignat-Lavaud, and S. Maffeis, "Discovering concrete attacks on website authorization by formal analysis," *Journal of Computer Security*, vol. 22, no. 4, pp. 601–657, 2014.

[43] G. Bai, J. Lei, G. Meng, S. S. Venkatraman, P. Saxena, J. Sun, Y. Liu, and J. S. Dong, "AUTHSCAN: automatic extraction of web authentication protocols from implementations," in *Network and Distributed System Security Symposium (NDSS)*, 2013.

[44] D. Fett, R. Küsters, and G. Schmitz, "An Expressive Model for the Web Infrastructure: Definition and Application to the BrowserID SSO System," in *35th IEEE Symposium on Security and Privacy (S&P 2014)*. IEEE Computer Society, 2014, pp. 673–688.

[45] K. Cohn-Gordon, C. Cremers, and L. Garratt, "On post-compromise security," in *IEEE Computer Security Foundations Symposium (CSF)*, 2016, pp. 164–178.

[46] K. Cohn-Gordon, C. Cremers, B. Dowling, L. Garratt, and D. Stebila, "A formal security analysis of the signal messaging protocol," in *IEEE European Symposium on Security and Privacy (Euro S&P)*, 2017.

[47] Y. Dodis, T. Ristenpart, J. Steinberger, and S. Tessaro, "To hash or not to hash again? (In)differentiability results for H^2 and HMAC," in *Advances in Cryptology (Crypto)*, 2012, pp. 348–366.

Appendix

A. ProScript to ProVerif Translation Samples

```
free io:channel.
type number. type function. type key. [...]
const string_63:bitstring [data]. (* WhisperMessageKeys *)
[...]
equation forall a:key, b:key;
  PS_crypto_DH25519(b, PS_crypto_DH25519(a, key_83)) =
  PS_crypto_DH25519(a, PS_crypto_DH25519(b, key_83)).
fun PS_crypto_AESCTREncrypt(key, iv, bitstring):bitstring.
reduc forall k:key, i:iv, m:bitstring;
     PS_crypto_AESCTRDecrypt(
  k, i, PS_crypto_AESCTREncrypt(k, i, m)) = m.
[...]
letfun fun_AKEResponse(me:me, them:them, msg:msg) =
let e = PS_crypto_random32Bytes(string_80) in let ge =
     PS_crypto_DH25519(e, key_83) in
let shared = fun_TDH0(keypair_get_priv(me_get_identity(me
     )), e, them_get_identity(them), msg_get_prekey(msg),
      msg_get_ephemeral(msg)) in
let recvKeys = fun_HKDF(shared, Type_key_construct(),
     string_56) in
let validSig = PS_crypto_checkED25519(them_get_identity(
     them), Type_key_toBitstring(msg_get_prekey(msg)),
     msg_get_prekeySig(msg)) in 0 [...]
free secMsg1:bitstring [private]. free secMsg2:bitstring
     [private].
query attacker(secMsg1). query attacker(secMsg2).
noninterf secMsg1. noninterf secMsg2.
event Send(key, key, bitstring). event Recv(key, key,
     bitstring).
query a:key,b:key,m:bitstring; event(Recv(a, b, m)) ⟹
     event(Send(a, b, m)).
[...]
let Alice(me:me, them:them) =
  let aStartSession = fun_startSession(me, them) in
  let them = sendoutput_get_them(aStartSession) in
  out(io, sendoutput_get_output(aStartSession));
  in(io, bAcceptSession:msg);
  let aAcceptSession = fun_recv(me, them, bAcceptSession)
       in
  let them = recvoutput_get_them(aAcceptSession) in
  let encMsg1 = fun_send(them, secMsg1) in
  let them = sendoutput_get_them(encMsg1) in
  event Send(keypair_get_pub(me_get_identity(me)),
       them_get_identity(them), secMsg1);
  out(io, sendoutput_get_output(encMsg1));
  phase 1; out(io, keypair_get_priv(me_get_identity(me)))
       ; in(io, encMsg2:msg);
  let decMsg2 = fun_recv(me, them, encMsg2) in
  if (msg_get_valid(recvoutput_get_output(decMsg2)) =
       true) then (
    let msg2 = recvoutput_get_plaintext(decMsg2) in
    event Recv(them_get_identity(them), keypair_get_pub(
         me_get_identity(me)), msg2)
  ).
let Bob(me:me, them:them) = [...]
let AliceM(me:me, them:them) = [...]
let BobM(me:me, them:them) = [...]
process
let alice = fun_newIdentity() in let bob =
     fun_newIdentity() in let mallory = fun_newIdentity()
      in
out(io, mallory);
let bobsAlice = [...] in let alicesMallory = [...] in let
      bobsMallory = [...]
(Alice(alice, alicesBob) | Bob(bob, bobsAlice) | AliceM(
     alice, alicesMallory) | BobM(bob, bobsMallory))
```

Figure 8: Extracted model with types, equations between primitives and human-readable protocol functionality.

```
const Type_key = {
  construct: function() {
    return [0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00
         ,
         0x00, 0x00, 0x00, 0x00. 0x00, 0x00, 0x00, 0x00,
         0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00,
         0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00, 0x00]
  },
  assert: function(a) {
    var i = 0; for (i = 0; i < 32; i++) {
      a[i&31] & 0x00}; return a
  },
  [...]
}
[...]
const RATCHET = {
  deriveSendKeys: function(them, newEphemeralPriv) {
    const kShared = ProScript.crypto.DH25519(
         newEphemeralPriv, them.ephemeral)
    const sendKeys = UTIL.HKDF(kShared, them.recvKeys[0],
         'WhisperRatchet')
    const kKeys = UTIL.HKDF(
      ProScript.crypto.HMACSHA256(sendKeys[1], '1'),
      Type_key.construct(), 'WhisperMessageKeys'
    )
    return {
      sendKeys: sendKeys, recvKeys: them.recvKeys,
      kENC: kKeys[0], kMAC: kKeys[1]
    }
  },
  deriveRecvKeys: function(them, newEphemeralPub) { [...]
       }
}
[...]
const TextSecure = {
  newIdentity: function() { [...] },
  startSession: function(me, them) {
    var me = Type_me.assert(me)
    var them = Type_them.assert(them)
    return {
      them: them,
      output: {
        status: 1,
        valid: true,
        prekey: Type_key.construct(),
        ephemeral: Type_key.construct(),
        [...]
      }
    }
  },
  acceptSession: function(me, them, msg) { [...] },
  send: function(them, plaintext) { [...] },
  recv: function(me, them, msg) {
    var me = Type_me.assert(me)
    var them = Type_them.assert(them)
    var msg = Type_msg.assert(msg)
    const themMsg = {them: them, msg: msg}
    if ((msg.status === 2) && (them.status === 0)) {
      return recvHandlers.AKEResponse(me, them, msg)
    }
    else if ((msg.status === 3) && (them.status === 2)) {
      return recvHandlers.message(recvHandlers.completeAKE
           (me, them, msg))
    }
    else if ((msg.status === 3) && (them.status === 3)) {
      return recvHandlers.message(themMsg)
    }
    else { return {
       them: Type_them.construct(), output: Type_msg.
            construct(), plaintext: ''
    }}
  }
}
```

Figure 9: SP functionality is written in ProScript's idiomatic style which employs JavaScript's purely functional programming language features.

B. ProScript Operational Semantics

Notation: Sorts and Constants.

$\mathtt{x} \in \mathcal{X}^{\mathrm{U}}$	User variable names.
$\{@proto, @this, @scope, @body\} \in \mathcal{X}^{\mathrm{I}}$	Internal variables.
$x \in \mathcal{X} \triangleq \mathcal{X}^{\mathrm{I}} * \mathcal{X}^{\mathrm{U}}$	Variable names.
$l \in \mathcal{L}_{\mathtt{null}} \triangleq \mathcal{L} \cup \{\mathtt{null}\}$	Locations.
l_{op}	`Object.prototype`.
l_g	Global object.
$L \in \mathcal{C} \triangleq \mathcal{L}^*$	Scope chains.
$\mathtt{n} \in \mathcal{N}$	Numbers.
$\mathtt{s} \in \mathcal{S}$	Strings.
$\mathtt{undefined} \in \mathcal{U}$	Undefined.
$\mathtt{e} \in \mathcal{E}$	Expressions.
$l{\cdot}\mathtt{x} \in \mathcal{R}$	References.
$\lambda \mathtt{x}{\cdot}\mathtt{e} \in \mathcal{F}$	Function code.
$\mathtt{p} \in \mathcal{P} \triangleq \mathcal{X}^{\mathrm{U}} \cup \mathcal{S}$	Property names.
$\mathtt{v} \in \mathcal{V}^{\mathrm{U}} \triangleq \mathcal{N} \cup \mathcal{S} \cup \mathcal{U} \cup \{\mathtt{null}\}$	User values.
$v \in \mathcal{V} \triangleq \mathcal{V}^{\mathrm{U}} \cup \mathcal{L}_{\mathtt{null}} \cup \mathcal{C} \cup \mathcal{F}$	Semantic values.
$H \in \mathcal{L} \times \mathcal{X} \to \mathcal{V}$	Heaps.

Notation: Functions and Judgements.

$\sigma(H, L, x)$	Find defining scope.
$\pi(H, L, x)$	Prototype resolution.
$\gamma(H, r)$	Dereferencing values.
$\mathsf{True}(E)$	E is true.
$\mathsf{False}(E)$	E is false.
$\mathsf{obj}(l, l')$	Empty object at l.
$\mathsf{fun}(l, L, \mathtt{x}, \mathtt{e}, l')$	New function at l.
$\mathsf{act}(l, \mathtt{x}, v, l', \mathtt{e}, l'')$	Activation object template.
$\mathsf{defs}(\mathtt{x}, l, \mathtt{e})$	Allocate local variable.

Scope Resolution: $\sigma(H, L, x)$.

$$\sigma(H, [], x) \triangleq \mathtt{null} \qquad \frac{\pi(H, l, x) \neq \mathtt{null}}{\sigma(H, l : L, x) \triangleq l}$$

$$\frac{\pi(H, l, x) = \mathtt{null}}{\sigma(H, l : L, x) \triangleq \sigma(H, L, x)}$$

Prototype Resolution: $\pi(H, L, x)$.

$$\pi(H, \mathtt{null}, x) \triangleq \mathtt{null} \qquad \frac{(l, x) \in \mathsf{dom}(H)}{\pi(H, l, x) \triangleq l}$$

$$\frac{(l, x) \notin \mathsf{dom}(H) \quad H(l, @proto) = l'}{\pi(H, l, x) \triangleq \pi(H, l', x)}$$

Dereferencing Values: $\gamma(H, r)$.

$$\frac{r \neq l{\cdot}\mathtt{x}}{\gamma(H, r) \triangleq r} \qquad \frac{\pi(H, l, \mathtt{x}) = \mathtt{null} \quad l \neq \mathtt{null}}{\gamma(H, l{\cdot}\mathtt{x}) \triangleq \mathtt{undefined}}$$

$$\frac{\pi(H, l, \mathtt{x}) = l' \quad l \neq \mathtt{null}}{\gamma(H, l{\cdot}\mathtt{x}) \triangleq H(l', \mathtt{x})}$$

Auxiliary Predicates.

$\mathsf{True}(E) \triangleq E \notin \{0, "", \mathtt{null}, \mathtt{undefined}\}$

$\mathsf{False}(E) \triangleq E \in \{0, "", \mathtt{null}, \mathtt{undefined}\}$

$\mathsf{obj}(l, l') \triangleq (l, @proto) \mapsto l'$

$\mathsf{fun}(F, Closure, Var, Body, Proto) \triangleq$
$(F, @scope) \mapsto Closure * (F, @body) \mapsto \lambda Var.Body *$
$(F, \mathtt{prototype}) \mapsto Proto * (F, @proto)$

$\mathsf{act}(l, \mathtt{x}, v, \mathtt{e}, l'') \triangleq l \mapsto$
$\{\mathtt{x} : v, @this : l'', @proto : null\} * \mathsf{defs}(\mathtt{x}, l, \mathtt{e})$

Local Variable Definition.

$$\frac{\mathtt{x} \neq \mathtt{y}}{\mathsf{defs}(\mathtt{x}, l, \mathtt{var\ y}) \triangleq (l, \mathtt{y}) \mapsto \mathtt{undefined}}$$

$\mathsf{defs}(\mathtt{x}, l, \mathtt{e1\ =\ e2}) \triangleq \mathsf{defs}(\mathtt{x}, l, \mathtt{e1})$

$\mathsf{defs}(\mathtt{x}, l, \mathtt{e1;e2}) \triangleq \mathsf{defs}(\mathtt{x}, l, \mathtt{e1}) \cup \mathsf{defs}(\mathtt{x}, l, \mathtt{e2})$

$\mathsf{defs}(\mathtt{x}, l, \mathtt{if(e1)\{e2\}\{e3\}}) \triangleq \mathsf{defs}(\mathtt{x}, l, \mathtt{e2}) \cup \mathsf{defs}(\mathtt{x}, l, \mathtt{e3})$

$$\frac{\textit{otherwise}}{\mathsf{defs}(\mathtt{x}, l, \mathtt{e}) \triangleq \mathtt{emp}}$$

Syntax of Terms: `v, e`.

$$\mathtt{v} \quad ::= (\ \mathtt{n} \mid \mathtt{m} \mid \mathtt{undefined} \mid \mathtt{null}\ )$$

$$\mathtt{e} \quad ::= \begin{pmatrix} \mathtt{e;e} \mid \mathtt{x} \mid \mathtt{v} \mid \mathtt{if(e)\ \{e\}\{e\}} \mid \mathtt{var\ x} \\ \mid \mathtt{const\ x} \mid \mathtt{e} \oplus \mathtt{e} \mid \mathtt{e.x} \mid \mathtt{e(e)} \mid \mathtt{x = e} \\ \mid \mathtt{function(x)e} \mid \mathtt{\{x_1:e_1 \ldots x_n:e_n\}} \mid \mathtt{e[e]} \end{pmatrix}$$

Operational Semantics: $H, L, \mathtt{e} \longrightarrow H', r$.

$$\text{Definition}\ \frac{}{H, L, \mathtt{(var \mid const)\ x} \longrightarrow H, \mathtt{undefined}}$$

$$\text{Value}\ \frac{}{H, L, \mathtt{v} \longrightarrow H, \mathtt{v}}$$

$$\text{MemberAccess}\ \frac{H, L, \mathtt{e} \longrightarrow H', l' \quad \gamma(H', l'{\cdot}\mathtt{x}) = v}{H, L, \mathtt{e.x} \longrightarrow H', v}$$

$$\text{Variable}\ \frac{\sigma(H, L, \mathtt{x}) = l' \quad \gamma(H, l'{\cdot}\mathtt{x}) = v}{H, L, \mathtt{x} \longrightarrow H, v}$$

$$\text{Object}\ \frac{H_0 = H * \mathsf{obj}(l', l_{op}) \quad \forall i \in 1..n.\ \begin{pmatrix} H_{i-1}, L, \mathtt{e}_i \longrightarrow H'_i, v_i \\ H_i = H'_i[(l', \mathtt{x}_i) \to v_i] \end{pmatrix}}{H, L, \{\mathtt{x}_1 : \mathtt{e}_1, \ldots, \mathtt{x}_n : \mathtt{e}_n\} \longrightarrow H_n, l'}$$

$$\text{Assignment}\ \frac{\sigma(H, L, \mathtt{x}) \longrightarrow l' \quad H, L, \mathtt{e} \longrightarrow H', v \quad H'' = H'[(l', \mathtt{x}) \to v]}{H, L, \mathtt{x} = \mathtt{e} \longrightarrow H'', v}$$

$$\text{Sequence}\ \frac{H, L, \mathtt{e}_1 \longrightarrow H', v \quad H', L, \mathtt{e}_2 \longrightarrow H'', v'}{H, L, \mathtt{e}_1\ ;\ \mathtt{e}_2 \longrightarrow H'', v'}$$

$$\text{BinaryOperators}\ \frac{H, L, \mathtt{e}_1 \longrightarrow H', \mathtt{v}_1 \quad H', L, \mathtt{e}_2 \longrightarrow H'', \mathtt{v}_2 \quad \mathtt{v}_1 \oplus \mathtt{v}_2 = v}{H, L, \mathtt{e}_1 \oplus \mathtt{e}_2 \longrightarrow H'', v}$$

$$\text{ComputedAccess}\ \frac{H, L, \mathtt{e}_1 \longrightarrow H_1, l' \quad l' \neq \mathtt{null} \quad H_1, L, \mathtt{e}_2 \longrightarrow H', \mathtt{p} \quad \gamma(H', l'\cdot\mathtt{p}) = v}{H, L, \mathtt{e}_1\,[\mathtt{e}_2] \longrightarrow H', v}$$

$$\text{Function}\ \frac{H' = H * \mathsf{obj}(l, l_{op}) * \mathsf{fun}(l', L, \mathtt{x}, \mathtt{e}, l)}{H, L, \mathtt{function}\,(\mathtt{x})\,\{\mathtt{e}\} \longrightarrow H', l'}$$

$$\text{FunctionCall}\ \frac{\begin{array}{c} H, L, \mathtt{e}_1 \longrightarrow H_1, l_1 \quad H_1(l_1, @body) = \lambda\ \mathtt{x}.\mathtt{e}_3 \quad H_1(l_1, @scope) = L' \\ H_1, L, \mathtt{e}_2 \longrightarrow H_2, v \quad H_3 = H_2 * \mathsf{act}(l, \mathtt{x}, v, \mathtt{e}_3, \bot) \quad H_3, l : L', \mathtt{e}_3 \longrightarrow H', v' \end{array}}{H, L, \mathtt{e}_1\,(\mathtt{e}_2) \longrightarrow H', v'}$$

$$\text{IfTrue}\ \frac{H, L, \mathtt{e}_1 \longrightarrow H', v \quad \mathsf{True}(v) \quad H', L, \mathtt{e}_2 \longrightarrow H'', v'}{H, L, \mathtt{if}\,(\mathtt{e}_1)\,\{\mathtt{e}_2\}\{\mathtt{e}_3\} \longrightarrow H'', v'}$$

$$\text{IfFalse}\ \frac{H, L, \mathtt{e}_1 \longrightarrow H', v \quad \mathsf{False}(v) \quad H', L, \mathtt{e}_3 \longrightarrow H'', v'}{H, L, \mathtt{if}\,(\mathtt{e}_1)\,\{\mathtt{e}_2\}\{\mathtt{e}_3\} \longrightarrow H'', v'}$$

C. CryptoVerif: Indifferentiability of HKDF

We start from the assumption that the compression function underlying SHA256 is a random oracle, and show that the functions

$$\begin{aligned} &x_1, x_2, x_3, x_4 \mapsto \mathsf{HKDF}(c_0 \| x_1 \| x_2 \| x_3 \| x_4, c_1, c_2) \\ &x_1, x_2, x_3 \mapsto \mathsf{HKDF}(c_0 \| x_1 \| x_2 \| x_3, c_1, c_2) \\ &x, y \mapsto \mathsf{HKDF}(x, y, c_2) \\ &x \mapsto \mathsf{HKDF}(x, c_1, c_4) \end{aligned}$$

where $c_1 = 0$ (256-bits), $c_2 =$ "WhisperRatchet", and $c_4 =$ "WhisperMessageKeys", can be considered as independent random oracles; more formally, we show that they are indifferentiable [32] from independent random oracles.

Definition 1 (Indifferentiability)**.** *A function F with oracle access to a random oracle H is (t_D, t_S, q, ϵ)-indifferentiable from a random oracle H' if there exists a simulator S such that for any distinguisher D*

$$|\Pr[D^{F,H} = 1] - \Pr[D^{H',S} = 1]| \leq \epsilon$$

The simulator S has oracle access to H' and runs in time t_S. The distinguisher D runs in time t_D and makes at most q queries.

In the game $G_0 = D^{F,H}$, the distinguisher interacts with the real function F and the random oracle H from which F is defined. In the game $G_1 = D^{H',S}$, the distinguisher interacts with a random oracle H' instead of F, and with a simulator S, which simulates the behavior of the random oracle H using calls to H'. Indifferentiability means that these two games are indistinguishable.

Theorem 4.4 in [47] shows that HMAC-SHA256 is then indifferentiable from a random oracle, provided the MAC keys are less than the block size of the hash function minus one, which is true here: the block size of SHA256 is 512 bits and the MAC keys are 256-bit long.

In the four calls to HKDF that we consider, we did not make explicit the length of the returned key material. This length is at most 512 bits, so we can consider that HKDF is defined by truncation of the following function HKDF_2 to the desired length:

$$\begin{aligned} &\mathsf{HKDF}_2(salt, key, info) = K_1 \| K_2 \text{ where} \\ &\quad prk = \mathsf{HMAC}(salt, key) \\ &\quad K_1 = \mathsf{HMAC}(prk, info \| 0x00) \\ &\quad K_2 = \mathsf{HMAC}(prk, K_1 \| info \| 0x01) \end{aligned} \tag{3}$$

and $\mathsf{HKDF}(key, salt, info)$ is a truncation of $\mathsf{HKDF}_2(salt, key, info)$. Much like for HMAC in [47], this function is not indifferentiable from a random oracle in general. Intuitively, the problem comes from a confusion between the first and the second (or third) call to HMAC, which makes it possible to generate prk by calling HKDF_2 rather than HMAC. In more detail, let

$$\begin{aligned} &prk \| \_ = \mathsf{HKDF}_2(s, k, i) \\ &salt = \mathsf{HMAC}(s, k) \\ &x = \mathsf{HMAC}(prk, info \| 0x00) \\ &x' \| \_ = \mathsf{HKDF}_2(salt, i \| 0x00, info) \end{aligned}$$

where the notation $x_1 \| x_2 = \mathsf{HKDF}_2(s, k, i)$ denotes that x_1 consists of the first 256 bits of $\mathsf{HKDF}_2(s, k, i)$ and x_2 its last 256 bits.

When HKDF_2 is defined from HMAC as above, we have $prk = \mathsf{HMAC}(prk', i \| 0x00)$ where $prk' = \mathsf{HMAC}(s, k) = salt$, so $prk = \mathsf{HMAC}(salt, i \| 0x00)$. Hence, $x' = \mathsf{HMAC}(prk, info \| 0) = x$. However, when HKDF_2 is a random oracle and HMAC is defined from HKDF_2, the simulator that computes HMAC sees what seems to be two unrelated calls to HMAC. (It is unable to see that prk is in

fact related to the previous call $salt = \mathsf{HMAC}(s, k)$: we have $prk\|\_ = \mathsf{HKDF}_2(s, k, i)$ but the simulator does not know which value of i it should use.) Therefore, the simulator can only return fresh random values for $salt$ and x, and $x \neq x'$ in general.

We can however recover the indifferentiability of HKDF_2 under the additional assumption that the three calls to HMAC use disjoint domains. Let $\mathcal{S}$, $\mathcal{K}$, and $\mathcal{I}$ be the sets of possible values of $salt$, key, and $info$ respectively, and $\mathcal{M}$ the set of 256-bit bitstrings, output of HMAC.

Lemma 1. *If $\mathcal{K} \cap (\mathcal{I}\|0x00 \cup \mathcal{M}\|\mathcal{I}\|0x01) = \emptyset$ and $\mathcal{S}$ consists of bitstrings of 256 bits, then HKDF_2 with domain $\mathcal{S} \times \mathcal{K} \times \mathcal{I}$ is (t_D, t_S, q, ϵ)-indifferentiable from a random oracle, where $\epsilon = \mathcal{O}(q^2/|\mathcal{M}|)$ and $t_S = \mathcal{O}(q^2)$, and $\mathcal{O}$ just hides small constants.*

Proof. Consider

- the game G_0 in which HMAC is a random oracle and HKDF_2 is defined from HMAC by (3), and
- the game G_1 in which HKDF_2 is a random oracle and HMAC is defined as follows.
  Let L be a list of pairs $((k, m), r)$ such that r is the result of a previous call to $\mathsf{HMAC}(k, m)$. The list L is initially empty.
  $\mathsf{HMAC}(k, m) =$
  1) if $((k, m), r) \in L$ for some r, then return r, else
  2) if $((k_0, m_0), k) \in L$ for some $k_0 \in \mathcal{S}$ and $m_0 \in \mathcal{K}$, and $m = info\|0x00$ for some $info \in \mathcal{I}$, then let $r\|\_ = \mathsf{HKDF}_2(k_0, m_0, info)$, else
  3) if $((k_0, m_0), k) \in L$ for some $k_0 \in \mathcal{S}$ and $m_0 \in \mathcal{K}$, and $m = k_1\|info\|0x01$ for some $k_1 \in \mathcal{M}$ and $info \in \mathcal{I}$, then let $k_1'\|k_2' = \mathsf{HKDF}_2(k_0, m_0, info)$; if $k_1' = k_1$, then $r = k_2'$;
  4) otherwise, let r be a fresh random element of $\mathcal{M}$;
  5) add $((k, m), r)$ to L;
  6) return r.

  We name *direct* oracle calls to HKDF_2 or HMAC calls that are done directly by the distinguisher, and *indirect* oracle calls the calls to HMAC done from inside HKDF_2 (in G_0) and the calls to HKDF_2 done from inside HMAC (in G_1).

Let us show that these two games are indistinguishable as long as, in G_0,

H1. HMAC never returns the same result for different arguments,

H2. no fresh result of HMAC is equal to the first argument of a previous call to HMAC,

H3. the distinguisher never calls $\mathsf{HMAC}(k, m)$ where $k = \mathsf{HMAC}(salt, key)$ has been called from inside HKDF_2 but not directly by the distinguisher,

H4. and $\mathsf{HMAC}(prk, info\|0x00)$ never returns a fresh k_1 such that $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ has been called (directly or indirectly) before,

and in G_1,

H5. there are no two elements $((k, m), r)$ and $((k', m'), r)$ in L with $(k, m) \neq (k', m')$,

H6. if the distinguisher calls $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ with $((salt, key), prk) \in L$ and $k_1\|\_ = \mathsf{HKDF}_2(salt, key, info)$, then $\mathsf{HKDF}_2(salt, key, info)$ has been called (directly or indirectly at step 2) before the call to $\mathsf{HMAC}(prk, k_1\|info\|0x01)$.

We have the following invariant:

P1. Given $salt, key$, there is at most one prk such that $((salt, key), prk) \in L$.

Indeed, when L contains such an element, calls to $\mathsf{HMAC}(salt, key)$ immediately return prk at step 1, and never add another element $((salt, key), prk')$ to L.

Case 1. Suppose the distinguisher makes a direct oracle call to HKDF_2 or HMAC with the same arguments as a previous direct call to the same oracle. Both G_0 and G_1 return the same result as in the previous call.

Case 2. Suppose the distinguisher makes a direct call to $\mathsf{HMAC}(k, m)$ with arguments that do not occur in a previous direct call to HMAC.

Case 2.a) In G_0, this HMAC call has already been done as $\mathsf{HMAC}(salt, key)$ from inside HKDF_2. In G_0, the result is $prk = \mathsf{HMAC}(salt, key)$, which is independent from previously returned values, so it looks like a fresh random value to the distinguisher. In G_1, we cannot have $m = info\|0x00$ nor $m = k_1\|info\|0x01$ because $m = key \in \mathcal{K}$ which is disjoint from $\mathcal{I}\|0x00$ and from $\mathcal{M}\|\mathcal{I}\|0x01$, so HMAC returns a fresh random value.

Case 2.b) In G_0, this HMAC call has already been done as $\mathsf{HMAC}(prk, info\|0x00)$ from inside $\mathsf{HKDF}_2(salt, key, info)$. Hence $\mathsf{HMAC}(k, m) = \mathsf{HMAC}(prk, info\|0x00)$ is the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ and $prk = \mathsf{HMAC}(salt, key)$. Since by H3, the distinguisher never calls $\mathsf{HMAC}(k, m)$ where $k = \mathsf{HMAC}(salt, key)$ has been called from inside HKDF_2 but not directly by the distinguisher, $\mathsf{HMAC}(salt, key)$ has been called directly by the distinguisher. In G_1, since $\mathsf{HMAC}(salt, key)$ has been called, $((salt, key), prk) \in L$, so $\mathsf{HMAC}(k, m) = \mathsf{HMAC}(prk, info\|0x00)$ returns the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ (step 2), as in G_0.

Case 2.c) In G_0, this HMAC call has already been done as $\mathsf{HMAC}(prk, K_1\|info\|0x01)$ from inside $\mathsf{HKDF}_2(salt, key, info)$. Hence $\mathsf{HMAC}(k, m) = \mathsf{HMAC}(prk, K_1\|info\|0x01)$ is the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$, $prk = \mathsf{HMAC}(salt, key)$, and $K_1 = \mathsf{HMAC}(prk, info\|0x00)$ is the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$. As above, $\mathsf{HMAC}(salt, key)$ has been called directly by the distinguisher. In G_1, since $\mathsf{HMAC}(salt, key)$ has been called, $((salt, key), prk) \in L$, so, since K_1 is the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$, $\mathsf{HMAC}(k, m) = \mathsf{HMAC}(prk, K_1\|info\|0x01)$ returns the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ (step 3), as in G_0.

Case 2.d) In G_0, this HMAC call has never been done, directly or indirectly. Hence, HMAC returns a fresh random value. In G_1, if $((salt, key), k) \in L$, then HMAC may return the first or last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$. However, since $\mathsf{HMAC}(k, m)$ has not been called from HKDF_2 in G_0, $\mathsf{HKDF}_2(salt, key, info)$ has not been called directly by the distinguisher, so the result of HMAC always looks like a fresh random value to the distinguisher.

Case 3. Suppose the distinguisher makes a direct call to $\mathsf{HKDF}_2(salt, key, info)$ with arguments that do not occur in a previous direct call to HKDF_2.

Case 3.a) In G_1, this call to HKDF_2 has already been done from HMAC. Hence $((salt, key), prk) \in L$ and $\mathsf{HMAC}(prk, info\|0x00)$ or $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ has been called. Since $((salt, key), prk) \in L$, $\mathsf{HMAC}(salt, key)$ has been called before the call to $\mathsf{HMAC}(prk, info\|0x00)$ or $\mathsf{HMAC}(prk, k_1\|info\|0x01)$, and it has returned prk.

Case 3.a) i) Suppose that $\mathsf{HMAC}(prk, info\|0x00)$ has been called and it returned k_1', and $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ has not been called. By step 2 of the definition of HMAC in G_1, since by H5, the only element of L of the form $(\_, prk)$ is $((salt, key), prk)$, $\mathsf{HMAC}(prk, info\|0x00)$ is the first 256 bits of a previous call to $\mathsf{HKDF}_2(salt, key, info)$. The current call to $\mathsf{HKDF}_2(salt, key, info)$ returns the same result, to its first 256 bits are $\mathsf{HMAC}(prk, info\|0x00)$. Its last 256 bits are independent from returned random values. Indeed, if a call to HMAC returns the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$, then this call occurs in step 3 of HMAC, and it is $\mathsf{HMAC}(prk', m)$ with $((salt, key), prk') \in L$, $m = k_1''\|info\|0x01$, and k_1'' is the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$. By P1, $prk' = prk$. We have $k_1' = k_1''$, so $\mathsf{HMAC}(prk', m)$ is $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$. But $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ has not been called by hypothesis, so no previous call to HMAC returns the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$. So the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ look like a fresh random value.

In G_0, the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ are also $\mathsf{HMAC}(prk, info\|0x00)$, where $prk = \mathsf{HMAC}(salt, key)$. Furthermore, the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ are independent of previously returned values. Indeed, $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ has not been called directly. Furthermore, it has not been called from previous calls to HKDF_2, because, if $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ had been called from $\mathsf{HKDF}_2(salt', key', info')$, then by $\mathcal{K} \cap (\mathcal{I}\|0x00 \cup \mathcal{M}\|\mathcal{I}\|0x01) = \emptyset$, this call would be the last of the three calls to HMAC in $\mathsf{HKDF}_2(salt', key', info')$, $prk = \mathsf{HMAC}(salt', key')$, and $info' = info$. Since by H1, HMAC never returns the same result for different arguments, this would imply $salt' = salt$ and $key' = key$, contradicting that $\mathsf{HKDF}_2(salt, key, info)$ has not been called before. Therefore, the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ look like a fresh random value.

Case 3.a) ii) Suppose that $\mathsf{HMAC}(prk, info\|0x00)$ has been called and it returned k_1', and $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ has been called. By definition of HMAC in G_1, $\mathsf{HMAC}(prk, info\|0x00)$ is the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ (step 2) and $\mathsf{HMAC}(prk, k_1'\|info\|0x01)$ is its last 256 bits (step 3), so $\mathsf{HKDF}_2(salt, key, info) = k_1'\|k_2'$ where $k_1' = \mathsf{HMAC}(prk, info\|0x00)$ and $k_2' = \mathsf{HMAC}(prk, k_1'\|info\|0x01)$. In G_0, we have the same property by definition of HKDF_2.

Case 3.a) iii) Otherwise, $\mathsf{HMAC}(prk, info\|0x00)$ has not been called.

If $\mathsf{HKDF}_2(salt, key, info)$ had been called from step 2 of HMAC, then we would have called $\mathsf{HMAC}(prk', m)$ with $((salt, key), prk') \in L$ and $m = info\|0x00$. Furthermore, by P1, $prk' = prk$, so we would have called $\mathsf{HMAC}(prk, info\|0x00)$. Contradiction. So $\mathsf{HKDF}_2(salt, key, info)$ has not been called from step 2 of HMAC.

Therefore, $\mathsf{HKDF}_2(salt, key, info)$ has been called from step 3 of HMAC. If $\mathsf{HKDF}_2(salt, key, info)$ had been called at step 3 of HMAC and its last 256 bits were returned, then the distinguisher would have called $\mathsf{HMAC}(prk', k_1'\|info\|0x01)$ with $((salt, key), prk') \in L$ and $k_1'\|\_ = \mathsf{HKDF}_2(salt, key, info)$. By H6, $\mathsf{HKDF}_2(salt, key, info)$ would have been called before, either directly (excluded by hypothesis) or indirectly at step 2. Then the distinguisher would have called $\mathsf{HMAC}(prk'', info\|0x00)$ with $((salt, key), prk') \in L$, so by P1, $prk'' = prk$, so this is excluded by hypothesis. Therefore, the last 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ were not returned at step 3.

We can then conclude that in G_1, the value of $\mathsf{HKDF}_2(salt, key, info)$ is independent from previously returned values, so it looks like a fresh random value.

In G_0, $\mathsf{HMAC}(salt, key)$ has been called directly and returned prk, $\mathsf{HMAC}(prk, info\|0x00)$ has not been called directly. If a previous call to $\mathsf{HKDF}_2(salt', key', info')$ called $\mathsf{HMAC}(prk, info\|0x00)$, then we would have $info' = info$ and $prk = \mathsf{HMAC}(salt', key')$. By H1, this would imply $salt' = salt$ and $key' = key$, so $\mathsf{HKDF}_2(salt, key, info)$ would have been called before, which is excluded by hypothesis. Therefore, $\mathsf{HMAC}(prk, info\|0x00)$ has not been called before, directly or indirectly. By H4, $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ has not been called before, with $k_1 = \mathsf{HMAC}(prk, info\|0x00)$. Therefore, $\mathsf{HMAC}(prk, info\|0x00)$ and $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ have not been called before, so their result is independent from previously returned values. Hence $\mathsf{HKDF}_2(salt, key, info)$ is independent from previously returned values, as in G_1.

Case 3.b) In G_1, this HKDF_2 call has never been done, directly or indirectly. Hence HKDF_2 returns a fresh random value. In G_0, the result is obtained from calls to HMAC. The distinguisher has not made these calls to HMAC directly calling $\mathsf{HMAC}(salt, key)$ first, because otherwise the simulator for HMAC in G_1 would have called $\mathsf{HKDF}_2(salt, key, info)$. Furthermore, it cannot call $\mathsf{HMAC}(salt, key)$ with result prk after calling $\mathsf{HMAC}(prk, info\|0x00)$ or $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ by H2. So the result of HKDF_2 is independent of the result of direct HMAC calls made by the distinguisher. Moreover, other calls to HKDF_2 did not generate the same last two calls to HMAC, because by H1, the first call to HMAC, $\mathsf{HMAC}(salt, key)$, never returns the same result for different arguments. So the result looks like a fresh random value to the distinguisher.

The previous proof shows that the games G_0 and G_1 are indistinguishable assuming the hypotheses H1–H6 hold. Let us bound the probability that they do not hold. Suppose that there are at most q (direct or indirect) queries to HMAC.

- The probability that H1 does not hold is at most the probability that among q random values in $\mathcal{M}$, two of them collide, so it is at most $q^2/|\mathcal{M}|$.
- The probability that H2 does not hold is at most the probability that among q random values in $\mathcal{M}$, one of them is equal to one among the q first arguments of HMAC queries, so it is also at most $q^2/|\mathcal{M}|$.
- When H3 does not hold, the distinguisher calls $\mathsf{HMAC}(k, m)$ for a value k that happens to be equal to $\mathsf{HMAC}(salt, key)$, which is independent of the values the distinguisher has seen, since $\mathsf{HMAC}(salt, key)$ has not been called directly by the distinguisher. There are at most q values $\mathsf{HMAC}(salt, key)$, and the distinguisher has q attempts, so the probability that H3 does not hold is at most $q^2/|\mathcal{M}|$.
- Similarly, when H4 does not hold, the fresh random value $\mathsf{HMAC}(prk, info\|0x00)$ collides with a previously fixed k_1. There are at most q values $\mathsf{HMAC}(prk, info\|0x00)$ and at most q values k_1, so the probability that H4 does not hold is at most $q^2/|\mathcal{M}|$.
- Let us show that, if the random values r chosen at step 4 are all distinct and distinct from first and second halves of HKDF_2 results used in HMAC, then H5 holds. The proof is by induction on the sequence of calls of HMAC. If $((k, m), r)$ is added to L and r comes from a result of HKDF_2 at step 2 or 3, then k determines k_0, m_0 uniquely by induction hypothesis, and m determines $info$ as well as which half of the result of HKDF_2 is r, hence r is uniquely determined from k, m, and distinct from elements chosen at step 4 by hypothesis. If $((k, m), r)$ is added to L and r is chosen at step 4, then r is always distinct from elements already in L by hypothesis. This concludes the proof of our claim.
  From this claim, we can easily see that the probability that H5 does not hold is at most $q^2/|\mathcal{M}|$.
- When H6 does not hold, the distinguisher calls $\mathsf{HMAC}(prk, k_1\|info\|0x01)$ and k_1 happens to be equal to the first 256 bits of $\mathsf{HKDF}_2(salt, key, info)$ which is independent from values returned to the distinguisher. So the probability that H6 does not hold is at most $q^2/|\mathcal{M}|$.

Hence, the probability that the distinguisher distinguishes G_0 from G_1 is at most $6q^2/|\mathcal{M}|$. □

The hypothesis of Lemma 1 is satisfied in our case because $\mathcal{K}$ consists of bitstrings of length 256 bits = 32 bytes or $3 \times 32 = 96$ bytes, $\mathcal{I}\|0x00$ consists of bitstrings of length at most 31 bytes and $\mathcal{M}\|\mathcal{I}\|0x01$ consists of bitstrings of length between 33 and 63 bytes.

Lemma 2. *If H is a random oracle, then the functions $H_1, \ldots, H_n$ defined as H on disjoint subsets $D_1, \ldots, D_n$ of the domain D of H are $(t_D, t_S, q, 0)$-indifferentiable from independent random oracles, where $t_S = \mathcal{O}(q)$ assuming one can determine in constant time to which subset D_i an element belongs.*

Proof. Consider

- the game G_0 in which H is a random oracle, and $H_i(x) = H(x)$ for each $x \in D_i$ and $i \leq n$, and
- the game G_1 in which $H_1, \ldots H_n$ are independent random oracles defined on $D_1, \ldots, D_n$ respectively, and $H(x) = H_i(x)$ if $x \in D_i$ for some $i \leq n$, and $H(x) = H_0(x)$ otherwise, where H_0 is a random oracle of domain $D \setminus (D_1 \cup \ldots \cup D_n)$.

It is easy to see that these two games are perfectly indistinguishable, which proves indifferentiability. □

By combining Lemmas 1 and 2, we obtain that

$$x_1, x_2, x_3, x_4 \mapsto \mathsf{HKDF}_2(0, c_0\|x_1\|x_2\|x_3\|x_4, c_2)$$
$$x_1, x_2, x_3 \mapsto \mathsf{HKDF}_2(0, c_0\|x_1\|x_2\|x_3, c_2)$$
$$x, y \mapsto \mathsf{HKDF}_2(x, y, c_2)$$
$$x \mapsto \mathsf{HKDF}_2(0, x, c_4)$$

are indifferentiable from independent random oracles. The domains are disjoint because different constants c_2 and c_4 are used and furthermore, the three cases that use c_2 differ by the length of their second argument ($4 \times 256 = 1024$ bits plus the length of c_0 for $x_1, x_2, x_3, x_4 \mapsto \mathsf{HKDF}_2(0, c_0\|x_1\|x_2\|x_3\|x_4, c_2)$, $3 \times 256 = 768$ bits plus the length of c_0 for $x_1, x_2, x_3 \mapsto \mathsf{HKDF}_2(0, c_0\|x_1\|x_2\|x_3, c_2)$, and 256 bits for $x, y \mapsto \mathsf{HKDF}_2(x, y, c_2)$).

Lemma 3. *If H is a random oracle that returns bitstrings of length l, then the truncation of H to length $l' < l$ is $(t_D, t_S, q, 0)$-indifferentiable from a random oracle, where $t_S = \mathcal{O}(q)$.*

Proof. Consider

- the game G_0 in which H is a random oracle, and $H'(x)$ is $H(x)$ truncated to length l', and
- the game G_1 in which H' is a random oracle that returns bitstrings of length l' and $H(x) = H'(x)\|H''(x)$ where H'' is a random oracle that returns bitstrings of length $l - l'$.

It is easy to see that these two games are perfectly indistinguishable, which proves indifferentiability. □

By combining Lemma 3 with the previous results, we conclude that

$$x_1, x_2, x_3, x_4 \mapsto \mathsf{HKDF}(c_0\|x_1\|x_2\|x_3\|x_4, c_1, c_2)$$
$$x_1, x_2, x_3 \mapsto \mathsf{HKDF}(c_0\|x_1\|x_2\|x_3, c_1, c_2)$$
$$x, y \mapsto \mathsf{HKDF}(x, y, c_2)$$
$$x \mapsto \mathsf{HKDF}(x, c_1, c_4)$$

are indifferentiable from independent random oracles.