2018 IEEE 31st Computer Security Foundations Symposium

Enforcing ideal-world leakage bounds in real-world secret sharing MPC frameworks

José Bacelar Almeida
INESC TEC and Universidade do Minho, Portugal

Manuel Barbosa
INESC TEC and FCUP Universidade do Porto, Portugal

Gilles Barthe
IMDEA Software Institute, Spain

Hugo Pacheco
INESC TEC and Universidade do Minho, Portugal

Vitor Pereira
INESC TEC and FCUP Universidade do Porto, Portugal

Bernardo Portela
INESC TEC and FCUP Universidade do Porto, Portugal

***Abstract*—We give a language-based security treatment of domain-specific languages and compilers for secure multi-party computation, a cryptographic paradigm that enables collaborative computation over encrypted data. Computations are specified in a core imperative language, as if they were intended to be executed by a trusted-third party, and formally verified against an information-flow policy modelling (an upper bound to) their leakage. This allows non-experts to assess the impact of performance-driven authorized disclosure of intermediate values.**

Specifications are then compiled to multi-party protocols. We formalize protocol security using (distributed) probabilistic information-flow and prove security-preserving compilation: protocols only leak what is allowed by the source policy. The proof exploits a natural but previously missing correspondence between simulation-based cryptographic proofs and (composable) probabilistic non-interference.

Finally, we extend our framework to justify leakage cancelling, a domain-specific optimization that allows to first write an efficient specification that fails to meet the allowed leakage upper-bound, and then apply a probabilistic pre-processing that brings leakage to the acceptable range.

I. Introduction

Secure multi-party computation (MPC) is a powerful cryptographic paradigm. MPC protocols allow two or more mutually distrusting parties to collaboratively compute over their private data, revealing nothing more than the result of the computation. MPC eliminates the need for delegating secure computations to a TTP (trusted third party), significantly reducing logistical and trust management problems, as well as security risks inherent to having a TTP as a single point of failure. As a consequence (and after two decades of sustained breakthroughs in its underlying technology) MPC is increasingly used for practical applications [1]–[3].

One key element for the practical success of MPC has been the emergence of domain-specific languages and compilers [4]–[12]. These MPC software stacks give (non-expert) programmers the ability to develop applications in traditional (sequential) programming languages, as if the computation was to be run by one TTP. These programs are then compiled to (probabilistic) protocols that realize the computation in a distributed, multi-party, setting.

To achieve efficient realizations, MPC programs tend to avoid computations that are expensive in a distributed setting, such as accessing arrays with secret indexes or securely branching based on secret values. A common approach to expose these constraints is via a standard information flow type system, with MPC-specific public control-flow restrictions (control-flow guards and array access expressions for imperative languages [13], or conditionals, fixpoint recursion [12], sum types and higher-order functions [6] for functional languages).[1]

The type system does not constrain the expressivity of the language, thanks to *declassify* statements, which turn an arbitrary expression to public. This suggests that its goal is *not* to enforce a secure information flow policy—a programmer is always free to declassify information—but to make programmers aware that the MPC application will perform some otherwise expensive computations publicly—as a performance optimization technique—and to ensure that data is consistently translated between public and secret semantic domains. Hence, while the type system is useful, it fails to capture rigorously *how much* information is released via declassification.

The lack of a rigorous mechanism for analyzing

[1] It is possible to express oblivious control-flow by computing all possible outcomes and algebraically selecting the correct result, but the overhead can be prohibitive in practical applications. Most often, this process can be performed as a compilation step (cf. [6], [12]), but we adopt the approach of [13] where programmers handle oblivious control-flow explicitly, which simplifies our presentation and formalization.

© 2018, José Bacelar Almeida. Under license to IEEE.
DOI 10.1109/CSF.2018.00017

IEEE computer society

leakage at source level is a serious hindrance for MPC technology, in particular because obtaining meaningful security guarantees has a significant impact on productivity. Even though high-level domain-specific source languages are tailored for non-experts, it is extremely hard to simultaneously achieve good performance, which implies declassifying intermediate results, and guarantee that leaked information is not harmful within a particular application, which usually calls for a MPC expert.

This paper demonstrates how to leverage language-based techniques to provide users of MPC domain-specific languages early and accurate feedback on the security of their programs. Technically, this is achieved in two steps: source-level analysis and secure compilation.

a) Source-level analysis: We propose an automated method for proving security of source programs. Our notion of security is expressed as a variant of non-interference, and states that inputs related by a leakage specification yield equal leakage, where leakage is modeled using an instrumented source-level semantics. Verification relies on relational program verification techniques, and is performed with minimal overhead. Indeed, we observe that MPC source programs, through information flow types and declassify statements, expose sufficient information to adapt a technique developed for analysing timing leaks in assembly code [14]. Our main contribution at this level is to adapt and extend this technique to deal with a real-world MPC programming language, and to demonstrate its application to proving meaningful (not trace-based) leakage upper-bounds. Using our tool for source-level analysis, a programmer that is not a MPC expert can prove a leakage upper-bound that can be matched to the security requirements of the application.

b) Secure compilation: We prove that low-level protocols do not leak more information than source programs from which they are generated. The central challenge here is to connect formally information flow-based notions of security for source programs and cryptographic simulation-based notions of security for protocols. Our solution is based on an alternative notion of protocol security, leveraging probabilistic information flow. We define a distributed probabilistic semantics that gives meaning to securely computing a functionality using a distributed protocol and introduce the notion of each party's view of the protocol. Our notion of protocol security states that parties executing the protocol correctly (a.k.a. honest-but-curious parties) cannot distinguish between two runs of the protocol on related inputs; precisely, the views—distributions over local execution traces—of each party are identical in the two runs.

We show that our security notion composes to justify *secure multiparty compilation* used in MPC software stacks: generating MPC protocols for arbitrary source programs by plugging simple atomic cryptographic components. Our main theorem states that, for any correctly typed program, source-level security is preserved as distributed information flow security of the compiled protocol. We conclude by proving that, for correct executions of the program, this implies the intended simulation-based notion of cryptographic security.

The challenge of secure compilation for MPC has been previously addressed by Mitchell et al. [12] and, indeed, our secure compilation theorem has a similar flavour. However, there are two main differences. i. We focus on secret sharing-based MPC, which allows us to give a unified probabilistic information-flow notion of protocol security that applies to both atomic and complex protocols. We prove that this property composes, greatly simplifying our secure compilation proof; we can work purely at the information-flow level, rather than reasoning inductively about the indistinguishability of distributions; and ii. We establish a natural but previously missing connection to standard security notions for information-theoretically secure MPC, by showing that our information-flow notion of protocol security is strong enough to imply the existence of a cryptographic simulator that requires only the leakage allowed by the source-level upper-bound to perfectly simulate real-world traces.

As an independent contribution related to performance, we leverage our framework to model *leakage cancelling*, a pattern usually performed by experts for optimizing MPC protocols in two steps: i. implement a specification p that leaks more than what is allowed, and then ii. use an efficient (oblivious) probabilistic preprocessing of inputs that renders leakage useless to an attacker. We give a sufficient condition (C) such that the following composition theorem holds: the sequential composition $p_0; p$ of a Ψ-secure program p_0 satisfying (C) and a Φ-secure program p is itself Φ-secure. Here, Φ-security means that two inputs satisfying a leakage specification Φ lead to identical leakage under program p.

Our main technical contributions are the following:

- an information flow-based definition of source-level leakage, and an automated method for proving that a program satisfies a leakage policy;
- an information flow-based definition of protocol security, and a proof that it entails the expected cryptographic notion of security;
- a proof (using a new technique) that compilation from programs to protocols preserves security;

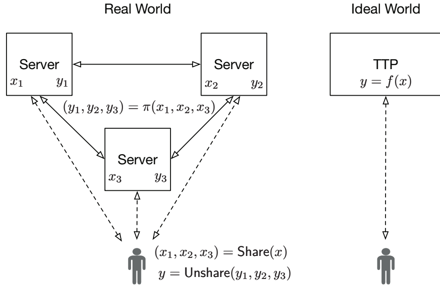

Figure 1: Real world versus ideal world.

- a formalization of leakage cancelling that attests its validity as a secure optimization technique;
- an implementation of our techniques for a real-world MPC language and an evaluation on challenging case studies from the literature.

All the proofs and more detailed technical discussions can be found in the accompanying technical report [15].

c) Limitations: Our language-based security framework is limited to passive security and private outputs. In the passive security model, honest-but-curious parties execute the protocol correctly, i.e., we do not consider adversaries that deviate from the protocol execution arbitrarily. Removing this assumption without compromising efficiency is an active area of research in cryptography, and extending our results to this setting is an important direction for further work. The standard security model for MPC protocols is universal composability, where attackers get to observe the raw protocol outputs. Our language-based security notion is slightly weaker than this, to practical gain, but it is well known [16] that a standard (and efficient) post-processing permits removing this caveat. We further discuss both limitations below. s

II. Overview and Motivating Examples

The family of MPC protocols that we study in this paper is based on secret sharing, a cryptographic primitive that permits splitting secret data between multiple parties, in such a way that accessing an incomplete subset of these *shares* reveals nothing about the secrets.[2]

Formally, a n-party secret sharing-scheme over a set S is defined by two algorithms: i. a probabilistic algorithm Share that takes a secret to be shared x and produces a distribution over n-tuples of shares $(x_1, \ldots, x_n) \in S^n$; and ii. a determistic algorithm $\mathsf{Unshare}$ that takes a tuple $(x_1, \ldots, x_n)$ and reconstructs the secret x. Correctness of secret sharing states that $\mathsf{Unshare}(\mathsf{Share}(x)) = x$ holds, for every x in S. For concreteness, we will consider an additive secret sharing scheme, where secrets are elements in a finite field $\mathbb{F}$. On input x, Share samples finite field elements $x_1, \ldots, x_n$ uniformly at random, conditioned on $x = x_1 + \ldots + x_n = \mathsf{Unshare}(x_1, \ldots, x_n)$.

The goal of MPC is then to perform (local or distributed) computations over shares, known as protocols, that are homomorphic to computations over the original secrets. This is naturally captured by the real-world vs. ideal world paradigm (see Figure 1 for a specialization to the 3-party case). On the right, the ideal world is emulated by the MPC platform: a user can offload secret data x to a TTP, that computes a function f over this data to obtain a result y. On the left, using a secret sharing-scheme, a user can offload shared data (x_1, x_2, x_3) to three servers in a secure way, and then obtain a secret-shared value (y_1, y_2, y_3) that results from the execution of a distributed protocol π between the three servers. We say that protocol π correctly implements f iff $\mathsf{Unshare}(y_1, y_2, y_3) = f(x)$, for every secret x and such that $(x_1, x_2, x_3) = \mathsf{Share}(x)$.

Designing MPC protocols requires trade-offs between efficiency and security. For instance, the following program, annotated with security types, avoids branching on secret values by using a declassify statement:[3]

```
secret minimum (secret xs) {
    secret min = x[0];
    for (i = 1; i < size(xs); i += 1)
        if declassify(x[i] < min) { min = x[i]; }
    return min; }
```

Even for this simple snippet, the first non-obvious question is how much information is leaked. Leakage at protocol level (messages exchanged among parties) is also rather different from source-level leakage (public values). The second question is whether a protocol π securely implements a program p, i.e. π does not leak more than p. In this paper we address both questions.

We provide a method for programmers to specify allowed leakage using annotations. For the above snippet, the programmer could declare that the function can leak comparisons between the vector elements. In the source annotation language proposed in this paper, this can be expressed as a pre-condition over an initial state xs:

```
forall uint i; 0<=i && i<size(xs)
    && 0<=j && j<size(xs) ==> public(xs[i] < xs[j])
```

Under the hood, this annotation is interpreted as a relational pre-condition on two initial states xs and xs':

[2]We will describe these protocols in the deployment setting typically adopted by platforms such as Sharemind, where there are one or more clients providing inputs and receiving outputs from the computation and a set of workers that carry out the computation over secret share data. Still, our results are also applicable to scenarios where input/output parties participate in the computation.

[3]Note that declassify statements can be *inferred*, but we follow the common practice of requiring that programmers manually insert them.

$$\begin{cases} \text{size(xs)} = \text{size(xs')} = k \\ \forall\, 0 \leq i, j < k.\text{xs}[i] < \text{xs}[j] \Leftrightarrow \text{xs}'[i] < \text{xs}'[j] \end{cases}$$

Our source-level verification can establish that this leakage bound is valid at source level using a notion of security that defines what source leakage is, and imposes that two executions of the above program on two related inputs states leak the same. A security preservation theorem then guarantees that the source-level bound is preserved by compilation to a distributed protocol.

Next, consider the following implementation of the partition operation for quick sort:

```
secret partition (secret xs, secret p) {
    for (i = 0; i < size (xs); i=i+1) {
        secret y = xs[i];
        if (declassify(y <= p)) {ls = snoc(ls,y);}
        else {rs = snoc(rs,y);}
    } return (ls,rs);  }
```

Our verification approach can show that the above leakage bound holds. Moreover, this leakage can be *cancelled* by probabilistic pre-processing. Intuitively, applying a random permutation to the quick sort input makes the sequence of comparisons look random, and useless to an attacker that does not know which permutation was applied. This yields a performance gain as random shuffling can be efficiently computed obliviously [17]. We leverage our formal framework to provide sound conditions for applying this optimization technique.

s

III. SOURCE-LEVEL LANGUAGE

a) Syntax: Our work considers SecreC [4], a commercial MPC language resembling C++, supporting high-level programming features such as procedures, templates or recursion, and used for writing secure applications in the Sharemind framework [13]. For our formal development, we will use a core imperative language extended with a declassify operator.[4] For clarity of presentation and w.l.o.g., we make a syntactic distinction between secure operations sop and public operations pop, and restrict the use of secure operations to top-level expressions. The syntax of programs appears in Figure 2.

b) Instrumented semantics: The semantics of our source language gives meaning to evaluating a MPC specification as if a TTP would be computing directly over the data, with full knowledge of secret and public variables. Despite being agnostic to security, this semantics is instrumented to construct a leakage trace including all branching conditions and all declassified values.

[4]This makes our results more widely applicable and helps distinguishing them from language features that are orthogonal to security analysis, but would complicate the formalism without additional insight.

$$\begin{array}{ll} p ::= \mathsf{skip} \mid p_1; p_2 \mid \mathsf{while}\, e\, \mathsf{do}\, p & lv ::= x \mid x[e] \\ \quad\mid lv := ae \mid \mathsf{if}\, e\, \mathsf{then}\, p_1\, \mathsf{else}\, p_2 & e ::= v \mid x \mid x[e] \\ ae ::= e \mid e_1\, \mathsf{sop}\, e_2 \mid \mathsf{declassify}\,(e) & \quad\mid e_1\, \mathsf{pop}\, e_2 \end{array}$$

Figure 2: Syntax of source-level language.

$$\frac{\langle p_1, m\rangle \to_l \langle \mathsf{skip}, m'\rangle \quad \langle p_2, m'\rangle \to_{l'} \langle p_2', m''\rangle}{\langle p_1; p_2, m\rangle \to_{l \cdot l'} \langle p_2', m''\rangle} \qquad \frac{\langle p_1, m\rangle \to_l \langle p_1', m'\rangle \quad p_1' \neq \mathsf{skip}}{\langle p_1; p_2, m\rangle \to_l \langle p_1'; p_2, m'\rangle}$$

$$\frac{p = \begin{cases} p_1 & \text{if } [\![e]\!](m) = \mathsf{tt} \\ p_2 & \text{if } [\![e]\!](m) = \mathsf{ff} \end{cases}}{\langle\, \mathsf{if}\, e\, \mathsf{then}\, p_1\, \mathsf{else}\, p_2, m\rangle \to_{[\![e]\!](m)} \langle p, m\rangle}$$

$$\frac{p' = \begin{cases} p;\, \mathsf{while}\, e\, \mathsf{do}\, p & \text{if } [\![e]\!](m) = \mathsf{tt} \\ \mathsf{skip} & \text{if } [\![e]\!](m) = \mathsf{ff} \end{cases}}{\langle\, \mathsf{while}\, e\, \mathsf{do}\, p, m\rangle \to_{[\![e]\!](m)} \langle p', m\rangle}$$

$$\frac{[\![e]\!](m) = v}{\langle lv := e, m\rangle \to_\varepsilon \langle\, \mathsf{skip}, m[lv \mapsto v]\rangle}$$

$$\frac{\mathsf{sop}\,([\![e_1]\!](m), [\![e_2]\!](m)) = v}{\langle lv := e_1\, \mathsf{sop}\, e_2, m\rangle \to_\varepsilon \langle\, \mathsf{skip}, m[lv \mapsto v]\rangle}$$

$$\frac{[\![e]\!](m) = v}{\langle lv := \mathsf{declassify}\,(e), m\rangle \to_v \langle\, \mathsf{skip}, m[lv \mapsto v]\rangle}$$

Figure 3: Source-level instrumented semantics.

The semantics is defined using the standard notion of transitions between configurations (Figure 3). A configuration $\langle p, m\rangle$ denotes a program p to be executed under a memory m. A memory m maps variables x or array elements $x[i]$ to values v. The evaluation of expression e under memory m is written $[\![e]\!](m)$. A transition from configuration c to c' is denoted by $\langle c\rangle \to_l \langle c'\rangle$. An execution of a program is then a sequence of configurations. Configuration $\langle p, m\rangle$ terminates in m' with leakage l, written $\langle p, m\rangle \Downarrow_l m'$, if $\langle p, m\rangle \to_l^* \langle\, \mathsf{skip}, m'\rangle$, where $\to^*$ forms the reflexive transitive closure of $\to$ and leakage is concatenated into a leakage trace.

Source-level semantics leaves as undefined the meaning of unsafe programs. A program is *safe* when its semantics is defined for every initial state. In particular, this entails that the program terminates on all inputs. This property can be checked using standard verification techniques supported by our tool mentioned further in the paper.

c) Source-level security: Our notion of source-level security is an information flow policy that sets an upper bound on the leakage of a secure program.

Definition 1 (Source-level security)**.** *Let* Φ *be relations over memories. A program* p *is* Φ-secure *whenever:*

$$\left(\begin{array}{l} \langle p, x_1\rangle \Downarrow_{l_1} y_1 \\ \langle p, x_2\rangle \Downarrow_{l_2} y_2 \end{array}\right) \Rightarrow \Phi(x_1, x_2) \Rightarrow l_1 = l_2$$

$$
\frac{}{\Gamma \vdash \mathsf{skip}} \qquad \frac{\Gamma \vdash p_1 \quad \Gamma \vdash p_2}{\Gamma \vdash p_1 ; p_2} \qquad \frac{\Gamma \vdash_e e : \mathsf{public} \quad \Gamma \vdash p}{\Gamma \vdash \mathsf{while}\, e\, \mathsf{do}\, p}
$$

$$
\frac{\Gamma \vdash_{lv} lv : s \quad \Gamma \vdash_{ae} ae : s' \quad s' \sqsubseteq s}{\Gamma \vdash lv := ae} \qquad \frac{\Gamma \vdash_e e : \mathsf{public} \quad \Gamma \vdash p_1 \quad \Gamma \vdash p_2}{\Gamma \vdash \mathsf{if}\, e\, \mathsf{then}\, p_1\, \mathsf{else}\, p_2}
$$

$$
\frac{\Gamma(x) = s}{\Gamma \vdash_{lv} x : s} \qquad \frac{\Gamma(x) = s \quad \Gamma \vdash_e e : \mathsf{public}}{\Gamma \vdash_{lv} x[e] : s}
$$

$$
\frac{}{\Gamma \vdash_e v : \mathsf{public}} \qquad \frac{\Gamma \vdash_{lv} lv : s}{\Gamma \vdash_e lv : s} \qquad \frac{\Gamma \vdash_e e_1 : \mathsf{public} \quad \Gamma \vdash_e e_2 : \mathsf{public}}{\Gamma \vdash_e e_1\, \mathsf{pop}\, e_2 : \mathsf{public}}
$$

$$
\frac{\Gamma \vdash_e e : s \quad s \sqsubseteq s'}{\Gamma \vdash_e e : s'} \qquad \frac{\Gamma \vdash_e e : \mathsf{private}}{\Gamma \vdash_{ae} \mathsf{declassify}\,(e) : \mathsf{public}}
$$

$$
\frac{\Gamma \vdash_e e : s}{\Gamma \vdash_{ae} e : s} \qquad \frac{\Gamma \vdash_e e_1 : \mathsf{private} \quad \Gamma \vdash_e e_2 : \mathsf{private}}{\Gamma \vdash_{ae} e_1\, \mathsf{sop}\, e_2 : \mathsf{private}}
$$

Figure 4: Type system of our source language.

Note that every program is secure w.r.t. full leakage, i.e., the equality relation, since the instrumented semantics is deterministic. Further, we can assume w.l.o.g. that Φ is an equivalence relation, as Φ-security and Φ^*-security coincide—as usual, Φ^* denotes the reflexive, symmetric and transitive closure of Φ. Finally, note that every function ℓ mapping states to an arbitrary type L induces a leakage relation $\Phi_\ell(x, y)$ defined as $\ell(x) = \ell(y)$.

The definition is an instance of observational non-interference, where it is required that observation traces l_1 and l_2 (rather than program outputs) are equivalent under related program inputs. An implication of source-level security is that, when dealing with Φ-equivalent inputs, two executions of a source-secure program are guaranteed to run in lock-step, i.e., all Φ-equivalent inputs have identical control flow. This is because all branching conditions are included in the leakage traces.

IV. Source-level security verification

a) Type system: As we have seen in Section II, MPC languages have an intrinsic notion of security domains, forming a security lattice $\mathcal{L}$ satisfying the ordering $\mathsf{public} \sqsubseteq \mathsf{private}$. We formalize a type system for our source-level language (Figure 4) that statically assigns security labels to intermediate variables. A security environment $\Gamma : V \to \mathcal{L}$ defines a mapping from variables to security labels, and $\Gamma(x)$ denotes the security label of variable x in environment Γ. We define typing judgments for programs p ($\Gamma \vdash p$), and auxiliarly for left-values lv ($\Gamma \vdash_{lv} lv : s$), expressions e ($\Gamma \vdash_e e : s$) and assignment expressions ae ($\Gamma \vdash_{ae} ae : s$) under security label s. Note that the type rules impose that branching conditions and array indices are public. As noted in the introduction, this is a design choice that we inherit from Sharemind [13].

Our type system is reminiscent of security type systems for information-flow with declassification [12], [18], which typically enforce trace-based notions of flow-insensitive non-interference or delimited release. Still, in this paper it only serves to syntactically restrict programs to a form—secret variables are only directly assigned to secret variables, or used as input to sop or declassify operations—compatible with our distributed semantic evaluation rules, where public values can be transparently used as secret ones, but the contrary is not true.[5]

b) Leakage analysis: A security type system can serve as a preliminary security analysis. Indeed, if declassify statements are *not* used (the program has *no* leakage), it can enforce source-level security for all leakage relations that guarantee equality of public inputs, but says little about security in the presence of leakage. Thus, we introduce a Security Hoare Logic (SHL) that provides a more powerful way to reason compositionally about source-level security w.r.t. a general leakage relation. We start by considering a security post-condition.

Definition 2 (Composable source-level security)**.** *A program p is (Φ, Ψ)-secure, written $\{\Phi\}\, p\, \{\Psi\}$, whenever:*

$$
\begin{pmatrix} \langle p, x_1 \rangle \Downarrow_{l_1} y_1 \\ \langle p, x_2 \rangle \Downarrow_{l_2} y_2 \end{pmatrix} \Rightarrow \Phi(x_1, x_2) \;\Rightarrow\; \Psi(y_1, y_2) \wedge l_1 = l_2
$$

The triple $\{\Phi\}\, p\, \{\Psi\}$ forms a security contract: the pre-condition Φ is an upper bound on the leakage of program p, and the post-condition Ψ evinces known leakage after running p. The rules of SHL are given in Figure 5, and are similar to those of standard Hoare Logic, with additional implications for leakage traces. Intuitively, it models the advice that a security verification system can give to a programmer who needs to justify leakage.

Formally, SHL has a relational interpretation consistent with source-level security, and reasons about pairs of executions of the same program running in lockstep.

Theorem 1. *$\vDash \{\Phi\}\, p\, \{\Psi\}$ is derivable iff $\{\Phi\}\, p\, \{\Psi\}$.*

A consequence of having public control-flow is that source-level security can be verified using deductive verification techniques over a self-composed program, with complexity in the same class as the original program.[6]

[5] Note that source-level security may allow the contrary if the secret values are publicly known from the leakage relation. This, in particular, is consistent to running a sop with public arguments; in practice, languages can offer a specialized sop relying on public information, e.g., instead of lifting a constant public value 3 and securely multiplying it by a secret variable y ($(*)\, 3\, y$), securely performing $(3*)\, y$.

[6] The verification techniques also support programs with secret control flow, but become less tractable. We could also apply our leakage verification after a secret-to-public control-flow compilation step.

$$\frac{}{\models \{\Phi\}\ \mathsf{skip}\ \{\Phi\}} \qquad \frac{\models \{\Phi\}\, p_1\, \{\Theta\} \quad \models \{\Theta\}\, p_2\, \{\Psi\}}{\models \{\Phi\}\, p_1 ; p_2\, \{\Psi\}}$$

$$\frac{\models \{\Phi \wedge e\}\, p_1\, \{\Psi\} \quad \models \{\Phi \wedge \neg e\}\, p_2\, \{\Psi\} \quad \Phi \Rightarrow \mathsf{public}\,(e)}{\models \{\Phi\}\ \mathsf{if}\ e\ \mathsf{then}\ p_1\ \mathsf{else}\ p_2\, \{\Psi\}} \qquad \frac{\Phi \Rightarrow \Phi' \quad \models \{\Phi'\}\, p\, \{\Psi'\} \quad \Psi' \Rightarrow \Psi}{\models \{\Phi\}\, p\, \{\Psi\}}$$

$$\frac{\models \{\Phi \wedge e\}\ \{\Phi\} \quad \Phi \Rightarrow \mathsf{public}\,(e)}{\models \{\Phi\}\ \mathsf{while}\ e\ \mathsf{do}\ p\, \{\Phi \wedge \neg e\}} \qquad \frac{\Phi = \Psi[e/lv]}{\models \{\Phi\}\, lv := e\, \{\Psi\}}$$

$$\frac{\Phi = \Psi[e_1\ \mathsf{sop}\ e_2/lv]}{\models \{\Phi\}\, lv := e_1\ \mathsf{sop}\ e_2\, \{\Psi\}} \qquad \frac{\Phi = \Psi[e/lv] \Rightarrow \mathsf{public}\,(e)}{\models \{\Phi\}\, lv := \mathsf{declassify}\,(e)\, \{\Psi\}}$$

$$\mathsf{public}\,(e)(x_1, x_2) \triangleq [\![e]\!](x_1) = [\![e]\!](x_2)$$

Figure 5: Inference system for Security Hoare Logic.

V. Low-level language

In this section we give a meaning to securely computing a source-level program using a distributed, secret sharing-based, cryptographic protocol. We will do this by providing a (low-level) distributed semantics for the specification language we introduced in the previous section. We start by introducing some notation.

a) Notation: Our low-level distributed semantics keeps a state of n separate maps $\mathsf{M} = (\mathsf{M}_1, \ldots, \mathsf{M}_n)$, each corresponding to the local state of a different party. Together, they satisfy the (informal) invariant that $(\mathsf{M}_1, \ldots, \mathsf{M}_n)$ encodes the state of the source-level computation. Each $\mathsf{M}_i : V \mapsto \{0, 1\} * S$ maps variables to pairs, where each stored value holds an encoding bit and a share. The encoding bit is used to identify values stored in shared form from a special encoding of public values. A variable v holding a shared value $\bar{x}$ will be stored in the n maps as $\mathsf{M}_i[v] = (0, x_i)$, for $1 \leq i \leq n$. To simplify the encoding of public values and w.l.o.g., we assume that n is odd. A variable v holding a public value c will be stored as $\mathsf{M}_i[v] = (1, c)$, for odd $1 \leq i \leq n$ and $\mathsf{M}_2[v] = (1, -c)$ for the remaining parties.

This representation allows to locally reconstruct c without communication, and is consistent with its shared representation, as $c = (n/2+1)c - (n/2)c$. The decision to share public values was taken so that there would be a greater integration with MPC. The public values can thus always be used as shared (secret) values, but the converse is not true. We will use the notation $\|(b, a)\|_i$ to represent the decoding of a stored value:

$$\|(b, a)\|_i \ := \ \mathsf{if}\ b = 1 \wedge i \bmod 2 = 0\ \mathsf{then}\ (-a)\ \mathsf{else}\ a\,.$$

The low-level distributed semantics assumes the existence of basic cryptographic protocols for all secure operators sop in the source language. Each of these is a n-party protocol π, whose (distributed) execution is denoted $(\bar{y}, t, c) \twoheadleftarrow \pi_{\mathsf{sop}}\,(\bar{x}, \bar{x}')$, as short-hand for: i. sampling random coins $c = (c_1, \ldots, c_n)$ in the appropriate spaces, ii. running the n parties on inputs $((x_1, x'1), \ldots, (x_n, x'n)) = (\bar{x}, \bar{x}')$ and random coins, iii. recording the interaction between parties in trace t, and iv. collecting outputs $(y_1, \ldots, y_n) = \bar{y}$. We will use t_i, for $1 \leq i \leq n$ to denote the part of the communications trace that is within the view of each party (both sent and received messages). We note that the local behaviour of each participant i is fully determined by its local input shares x_i, x'_i its random coins c_i and its local trace t_i. We can therefore meaningfully refer to the recomputing of a local output as $y_i \leftarrow \pi_i(x_i, x'_i, t_i, c_i)$.

$\pi_{\mathsf{declassify}}\,(u_1, \ldots, u_n; c_1, \ldots, c_n)$:
- For $i = 1$ to $n-1$: P_i sends c_i to P_{i+1}; P_n sends c_n to P_1
- P_1 computes $u'_1 \leftarrow u_1 + c_n - c_1$, broadcasts u'_1
- For $i = 2$ to n: P_i computes $u'_i \leftarrow u_i + c_{i-1} - c_i$, broadcasts u'_i
- For $i = 1$ to n, s.t. i is odd:
  - P_i computes $u = \mathsf{Unshare}(\bar{u}')$, locally returns u
- For $i = 1$ to n, s.t. i is even:
  - P_i computes $u = \mathsf{Unshare}(\bar{u}')$, locally returns $(-u)$

Figure 6: Declassify protocol.

For illustrating the two semantic domains, we explicitly define a $\pi_{\mathsf{declassify}}$ protocol (Figure 6), that moves secret shared distributed values to public local values.

b) Distributed semantics: The distributed semantics for the SecreC language is presented in Figure 7. It relies on the following expression evaluation rules.

$$\begin{aligned}
[\![v]\!](\mathsf{M}_i) &= \mathsf{M}_i[v] \\
[\![x]\!](\mathsf{M}_i) &= \mathsf{M}_i(x) \\
[\![x[e]]\!](\mathsf{M}_i) &= \mathsf{M}_i(x)[\|[\![e]\!](\mathsf{M}_i)\|_i] \\
[\![e_1\ \mathsf{pop}\ e_2]\!](\mathsf{M}_i) &= (1, v)
\end{aligned}$$

$$\begin{aligned}
\text{where}\quad & v_1 = \|[\![e_1]\!](\mathsf{M}_i)\|_i \\
& v_2 = \|[\![e_2]\!](\mathsf{M}_i)\|_i \\
& \|(1, v)\|_i = \mathsf{pop}\,(v_1, v_2)
\end{aligned}$$

We present the semantics from a local evaluation perspective, as transitions do not require interaction between the parties, except for the evaluation rules of sop or $\mathsf{declassify}$ operators that have explicit communication:

$$\begin{aligned}
[\![e_1\ \mathsf{sop}\ e_2]\!]_{ae}((\mathsf{M}_1, \ldots, \mathsf{M}_n)) &= (\bar{v}', t, c) \\
\text{where}\quad & \forall 1 \leq i \leq n,\ [\![e_1]\!](\mathsf{M}_i) = (\cdot, v_{1,i}) \\
& \forall 1 \leq i \leq n,\ [\![e_2]\!](\mathsf{M}_i) = (\cdot, v_{2,i}) \\
& (\bar{v}', t, c) \twoheadleftarrow \pi_{\mathsf{sop}}\,(\bar{v}_1, \bar{v}_2) \\
[\![\mathsf{declassify}\,(e)]\!]_{ae}((\mathsf{M}_1, \ldots, \mathsf{M}_n)) &= (\bar{v}', t, c) \\
\text{where}\quad & \forall i,\ [\![e]\!](\mathsf{M}_i) = (\cdot, v_i) \\
& (\bar{v}', t, c) \twoheadleftarrow \pi_{\mathsf{declassify}}\,(\bar{v})
\end{aligned}$$

$$\frac{}{\langle \mathsf{skip}\,; p, \mathsf{M}_i\rangle \Rrightarrow \langle p, \mathsf{M}_i\rangle} \qquad \frac{\langle p_1, \mathsf{M}_i\rangle \Rrightarrow \langle p_1', \mathsf{M}_i'\rangle}{\langle p_1; p_2, \mathsf{M}_i\rangle \Rrightarrow \langle p_1'; p_2, \mathsf{M}_i'\rangle} \qquad \frac{[\![e]\!](\mathsf{M}_i) = (b, v)}{\langle lv := e, \mathsf{M}_i\rangle \Rrightarrow \langle \mathsf{skip}\,, \mathsf{M}_i[lv \mapsto (b, v)]\rangle}$$

$$\frac{p_j = \begin{cases} p_1 & \text{if } \|[\![e]\!](\mathsf{M}_i)\|_i = \mathsf{tt} \\ p_2 & \text{if } \|[\![e]\!](\mathsf{M}_i)\|_i = \mathsf{ff} \end{cases}}{\langle \text{ if } e \text{ then } p_1 \text{ else } p_2, \mathsf{M}_i\rangle \Rrightarrow \langle p_j, \mathsf{M}_i\rangle} \qquad \frac{p' = \begin{cases} p;\ \text{while } e \text{ do } p & \|[\![e]\!](\mathsf{M}_i)\|_i = \mathsf{tt} \\ \mathsf{skip} & \|[\![e]\!](\mathsf{M}_i)\|_i = \mathsf{ff} \end{cases}}{\langle \text{ while } e \text{ do } p, \mathsf{M}_i\rangle \Rrightarrow \langle p', \mathsf{M}_i\rangle}$$

$$\frac{[\![e_1 \mathsf{\ sop\ } e_2]\!]_{ae}(\mathsf{M}) = (\bar{v}', t, c)}{\langle lv := e_1 \mathsf{\ sop\ } e_2, \mathsf{M}\rangle \Rrightarrow_{t,c} \langle \mathsf{skip}\,, \forall i : \mathsf{M}_i[lv \mapsto (0, v_i')]\rangle} \qquad \frac{[\![\, \mathsf{declassify}\ (e)]\!]_{ae}(\mathsf{M}) = (\bar{v}', t, c) \quad b_{1,3,5\ldots} = 0 \quad b_{2,4,6\ldots} = 1}{\langle lv := \mathsf{declassify}\ (e), \mathsf{M}\rangle \Rrightarrow_{t,c} \langle \mathsf{skip}\,, \forall i : \mathsf{M}_i[lv \mapsto (b_i, v_i')]\rangle}$$

Figure 7: Low-level distributed semantics.

The pre-requisites in these two transition rules will block the progress of all local evaluations until all parties synchronously execute them. Furthermore, only these rules contribute to the global execution trace.

As before, $\Rrightarrow^*$ forms the reflexive transitive closure of $\Rrightarrow$. The execution trace for the distributed evaluation of the program is the concatenation of rule traces. An execution of a program is then a sequence of distributed configurations. Configuration $\langle p, (\mathsf{M}_1, \ldots, \mathsf{M}_n)\rangle$ terminates execution in configuration $(\mathsf{M}_1', \ldots, \mathsf{M}_n')$ with trace (t, c), written $\langle p, (\mathsf{M}_1, \ldots, \mathsf{M}_n)\rangle \Downarrow_{(t,c)} (\mathsf{M}_1', \ldots, \mathsf{M}_n')$, if

$$\forall 1 \le i \le n : \langle p, (\mathsf{M}_i)\rangle \Rrightarrow^*_{(t,c)} \langle (\, \mathsf{skip}\,), (\mathsf{M}_i')\rangle\,.$$

We slightly abuse notation and refer to the distributed evaluation of secure operators sop and declassify as

$$\langle\, \mathsf{sop}\,, \bar{x}\rangle \Downarrow_{(t,c)} \bar{x}' \qquad \langle\, \mathsf{declassify}\,, \bar{x}\rangle \Downarrow_{(t,c)} \bar{x}'\,.$$

We have intentionally written our distributed semantics so that it can be seen as a high-level cryptographic protocol that relies on lower level ones to evaluate a program p. Put differently, our distributed semantics describes a compiler that takes a high level MPC specification and produces a composite protocol π_p. We will prove strong cryptographic security properties for this protocol in the style of certified compilation: if p is source-level secure, then π_p will guarantee that this security is translated into standard cryptographic security guarantees.

c) Low-level correctness and security: To reason about the guarantees provided by our MPC software stack, we introduce correctness and security notions for low-level evaluations. The definitions apply to high-level cryptographic protocols π_p and also to low-level cryptographic protocols π_{sop} and $\pi_{\mathsf{declassify}}$.

Intuitively, correctness states that whatever behaviours are observable at the source level will be also observable at target level, modulo the sharing relation.

Definition 3 (Low-level correctness)**.** *A protocol π is low correct for specification $s \in \{p, \mathsf{sop}\,, \mathsf{declassify}\,\}$ if, for all sharings $\bar{x}$, we have*

$$\langle s, \mathrm{Unshare}(\bar{x})\rangle \Downarrow \mathrm{Unshare}(\bar{y}) \Rightarrow \langle \pi, \bar{x}\rangle \Downarrow_{t,c} \bar{y}$$

We define security by means of a probabilistic non-interference notion.

Definition 4 (Low-level security)**.** *Let Φ be a relation over unshared inputs. A protocol π is secure for Φ if, for all $1 \le i \le n$ and all sharings $\bar{x}$, $\bar{x}'$ such that $x = \mathrm{Unshare}(\bar{x})$ and $x' = \mathrm{Unshare}(\bar{x}')$ we have*

$$\Phi(x, x') \wedge \left(\begin{array}{c} \langle \pi, \bar{x}\rangle \Downarrow_{t,c} \bar{y} \\ \langle \pi, \bar{x'}\rangle \Downarrow_{t',c'} \bar{y'} \end{array} \right) \Rightarrow (t_i, c_i) = (t_i', c_i')$$

Informally, this definition states that messages exchanged (traces t) and randomness (coins c), as seen by each party throughout two distinct executions of the same protocol, will have identical distributions and thus leak no information about the (unshared) inputs in addition to that revealed by the leakage relation; two distributions are equal iff they assign the same probability to every element in their support.

Note that leakage relations are expressed over (unshared) values, and therefore low-secure protocols guarantee that no information about specific shares is revealed. This is extremely important when relating our security notion to cryptographic security definitions, where the attacker is able to see part of the shares.

d) Compositional reasoning: Compositionality is an important property of the correctness and security definitions for low-level distributed executions.

Lemma 1 (Low-level composability)**.** *Let protocols π_1, π_2 be correct w.r.t. programs p_1, p_2 and secure w.r.t. leakage relations Φ_1, Φ_2. Then, $\pi(\bar{x}) = \pi_2(\pi_1(\bar{x}))$ is correct for $p_2 \circ p_1$ and secure for $\Phi = \Phi_1 \wedge \Phi_2 \circ p_1$. Similarly, $\pi(\bar{x}) = (\pi_1(\bar{x}), \pi_2(\bar{x}))$ is correct for $p_2 \times p_1$ and secure for $\Phi_1 \wedge \Phi_2$.*

The proof of Lemma 1 can be found in [15].

VI. Security-preserving compilation

Our main theorem shows that our MPC software stack preserves source-level information-flow security to probabilistic distributed non-interference at the low-level. As a result, we will show in the next section that such systems guarantee security in the cryptographic sense.

Theorem 2 (Main Theorem)**.** *Assume by hypothesis that all π_{sop} are correct and secure for the leakage relation that accepts all inputs (nothing leaks). Let p be a program, such that $\Gamma \vdash p$ for some security context Γ, and let Φ be a leakage relation. Then, if p is source-level secure for Φ, we have that π_p is correct for p and secure for Φ.*

Proof. The proof proceeds in two steps that we formalize using two lemmas presented below. We first prove in Lemma 2 correctness of the low-level execution by relying on the composability of low-level correctness and the fact that the program type-checks (this guarantees that no secret-encoded value is used in a local public computation). We then prove in Lemma 3 that, for all MPC specifications p, the composability of low-security implies that the distributed execution will leak no more than the source-level traces. In other words, we get low-level security for the low-level leakage function that imposes equality of source-level traces (the leakage function is based on the source-level instrumented semantics). By transitivity, we can therefore conclude that π_p is Φ-low-secure, as source-level security guarantees source-level trace equality over inputs satisfying Φ. □

Lemma 2. *Assume by hypothesis that all π_{sop} are sop-low-correct. Let p be a program such that $\Gamma \vdash p$ for some security context Γ. Then π_p is correct for p.*

Lemma 3. *Assume by hypothesis that all π_{sop} are low-secure for the empty leakage relation. Let p be a program such that $\Gamma \vdash p$ for some security context Γ. Let also Φ_p be the leakage relation that imposes source-level leakage equality. Then, we have that π_p is Φ_p-secure.*

The proofs of Lemmas 2 and 3 can be found in [15].

VII. Cryptographic security

Cryptographic security of MPC protocols is typically defined using the simulation paradigm. Intuitively, the definition states that no attacker can distinguish a real world from an ideal world. In the real world, an attacker $\mathcal{A}$ interacts with the cryptographic protocol π directly, according to a set of rules that define an attack model In the ideal world, the parties relying on the protocol have access to an ideal functionality $\mathcal{F}$ representing a TTP.

game $\mathbf{Real}_{\Pi,\mathcal{A}}(\,)$:	**game** $\mathbf{Ideal}_{\mathcal{F}\langle p,\ell\rangle,\mathcal{S},\mathcal{A}}(\,)$:
$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1(\,)$	$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1(\,)$
	$x \leftarrow \mathrm{Unshare}(\bar{x})$
	$y \leftarrow p(x)$
$(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$	$l \leftarrow \ell(x)$
$y \leftarrow \mathrm{Unshare}(\bar{y})$	$(y_i, t, c) \twoheadleftarrow \mathcal{S}(i, l, x_i)$
Return $\mathcal{A}_2(\bar{x}, y, y_i, t_i, c_i, \mathsf{st})$	Return $\mathcal{A}_2(\bar{x}, y, y_i, t_i, c_i, \mathsf{st})$

Figure 8: Cryptographic privacy.

The ideal functionality also defines what information from the participant's inputs can be leaked to the attacker. Then, a simulator $\mathcal{S}$ must be able to fool the attacker into thinking it is actually in the real world based on this leakage, even when the attacker sees the outputs produced by the ideal functionality on inputs of its choosing. The existence of $\mathcal{S}$ shows that whatever the adversary sees in its attack can contain no more information about the inputs than what is specified by the functionality in the ideal world. Furthermore, since the attacker can observe the protocol outputs, and compare them to the ideal functionality output, it also implies that the protocol must be correct. The simulation-based definition that we adopt, and discuss later in this section, is called *privacy* [16]. Our notion of simulation-based security implies standard (computational) MPC security [19], since our simulators are polynomial time. This means that the resulting protocols can be directly used by cryptographers in combination with other constructions satisfying computational security in a compositional setting such as the Universal Composability framework.

We will consider a simple class of ideal functionalities. We denote such functionalities as $\mathcal{F}\langle p, \ell\rangle$, to indicate that they are parameterized by a source program p and a leakage function ℓ. The functionality specifies what a cryptographic protocol should achieve when executed on some shared initial state $\bar{x} = (\mathsf{M}_1, \ldots, \mathsf{M}_n)$:

i. produce a result y such that $y = p(\mathrm{Unshare}(\bar{x}))$, where the meaning of this evaluation is given by the source-level semantics (in cryptography all specifications are total, and so p must be safe);
ii. Leak at most $\ell(\mathrm{Unshare}(\bar{x}))$ information about the unshared input.

The security definition is given in Figure 8.

The attacker gets to pick the shared input and the identity of one party that will be corrupted.[7] This means that the attacker will know whatever this party knows, and still the protocol must leak only what the ideal

[7] The corruption of a single party is made for clarity of presentation. The same model and results can trivially be generalised to allow for an arbitrary set of corrupt parties under some adversarial threshold.

functionality specifies. In practice this ensures that, if there is an honest majority that does not collude to break the protocol, security is guaranteed.

In the real world the attacker observes the unshared protocol output and the part of the trace corresponding to the corrupted party. In the ideal world, the attacker sees the ideal functionality output and a simulated trace. We note that the ideal world simulator gets to see the input share of the corrupt party and the allowed leakage, and it must simulate the rest of the corrupt party's view.

Definition 5 (Cryptographic privacy)**.** *We say a protocol π is (ℓ, p)-private if there exists a probabilistic polynomial-time simulator $\mathcal{S}$ such that, for all adversaries $\mathcal{A}$, the following definition of advantage is* 0*:*

$$\mathbf{Adv}_{\Pi,\mathcal{F},\mathcal{A},\mathcal{S}} = \Pr[\mathsf{Real}_{\Pi,\mathcal{A}}] - \Pr[\mathsf{Ideal}_{\mathcal{F},\mathcal{S},\mathcal{A}}],$$

where games Real *and* Ideal *are described in Figure 8.*

We do not impose a bound on the attacker's computational power, meaning that we obtain information-theoretic security. The following theorem states that low-level security and correctness as defined in Section V imply cryptographic privacy, which allows us to rely only on language-based security techniques to reason about the security of our MPC software stack. Our result requires leakage functions to be efficiently invertible, in the sense that there exists a polynomial-time algorithm that, given l, computes some input x' such that $\ell(x') = l$. If this is not the case, then we get a weaker notion of security where the simulator must also be unbounded (the proof is the same except for this detail).

Theorem 3 (Cryptographic privacy from probabilistic non-interference)**.** *Let π be a protocol, p a safe program in our source language and ℓ a leakage function. Assume also that ℓ is efficently invertible.*[8] *Then, if π is low-correct for p and it is low-secure for leakage relation Φ_ℓ, we have that π is (ℓ, p)-private.*

A full game-based version of the proof can be found in Appendix A. Combined with the main theorem in the previous section, this result yields the following corollary.

Corollary 1 (Privacy-preserving compilation)**.** *Assume that all π_{sop} are correct and secure for the empty leakage relation. Let p be a safe program such that $\Gamma \vdash p$ for some security context Γ and let ℓ be a leakage function. Then, if p is Φ_ℓ-secure, we have that π_p is (p, ℓ)-private.*

[8] Requiring leakage functions to be efficiently invertible is not a significant caveat. This is true for the empty leakage function, and for all practical examples we have encountered, except if p is computing a cryptographic function for which efficient inverters are not known.

a) Relation to Universal Composability: The standard model for describing security of secure computation protocols is the Universal Composability (UC) framework [20]. The notion of privacy that we consider in this paper is weaker than UC-security, because the attacker does not see the raw output of the protocol in its shared form. However, this is not a significant limitation. Indeed, our results readily extend to a UC-realization of an arithmetic black box (ABB) [21] using standard techniques. ABBs are a common abstraction of secure computation applications when the goal is to design a system that performs several basic operations before producing an output [22]–[24]. Furthermore, the presented model considers static corruptions, while results in [16] consider adaptive corruptions. In [15], we discuss how our results can easily be extended to adaptive corruptions under the same assumptions.

VIII. Leakage Cancelling

The resolution of the security versus performance contention in MPC has led to interesting optimization techniques. One such technique consists of composing a program p that is secure for a leakage relation Φ with a probabilistic pre-processing step p_0, resulting in a (probabilistic, yet) functionally equivalent program p' that satisfies a weaker leakage relation Ψ. The requirement that Φ implies Ψ reflects the natural information-theoretic interpretation of relations, and ensures that the program p' leaks less information than p. A typical example is sorting: p leaks the length of the array and the sequence of the comparison, as modeled by the relation Φ_{cmp} defined in Section II, and we want p' to only leak the length of the array. Leakage cancelling can be achieved in this case by obliviously randomly shuffling the input array. We provide a rigorous justification of this technique.

We model probabilistic behaviors by extending the expression language with probabilistic operators. The instrumented semantics of programs is modified accordingly: $\langle p, x\rangle \Downarrow_{\tilde{l}} \tilde{y}$ now states that executing program p with initial memory x terminates with distribution $\tilde{y}$ on output memories and a distribution on leakage traces.[9] When programs have deterministic leakage, i.e., all their guards and declassified expressions do not depend on values computed by probabilistic operators, we write $\langle p, x\rangle \Downarrow_l \tilde{y}$ to state that executing program p with initial memory x terminates with distribution $\tilde{y}$ on output memories and leakage trace l.

[9] We implicitly assume that programs are safe, so that we consider distributions rather than sub-distributions, and leakage traces have bounded length. A more precise semantics would consider the joint distribution of $\tilde{y}$ and $\tilde{l}$, but this is not required for our purposes.

The following theorem provides sufficient conditions for leakage cancelling, stated in terms of a lifting of source-level security for arbitrary probabilistic programs.

Definition 6 (Probabilistic source-level security)**.** *A probabilistic program p is Φ-secure whenever:*

$$\Phi(x_1, x_2) \wedge \begin{pmatrix} \langle p, x_1 \rangle \Downarrow_{\tilde{l_1}} \tilde{y_1} \\ \langle p, x_2 \rangle \Downarrow_{\tilde{l_2}} \tilde{y_2} \end{pmatrix} \Rightarrow \tilde{l_1} = \tilde{l_2}$$

The following theorem is proved in [15].

Theorem 4 (Secure pre-processing)**.** *Let p_0 and p be Ψ-secure and Φ-secure programs, respectively, such that p_0 has deterministic leakage. Then, for every input states x_1, x_2 such that $\Psi(x_1, x_2)$, $\langle p_0, x_1 \rangle \Downarrow_{l_1} \tilde{y_1}$, $\langle p_0, x_2 \rangle \Downarrow_{l_2} \tilde{y_2}$ and all output states y:*

$$\Pr_{y_1 \leftarrow \tilde{y_1}}[\Phi(y_1, y)] = \Pr_{y_2 \leftarrow \tilde{y_2}}[\Phi(y_2, y)]$$

In particular, Theorem 4 holds for all pre-processing functions that yield uniformly distributed outputs. Technically, for every x, $\tilde{y}$, y_1 and y_2 [25]:

$$\langle p_0, x \rangle \Downarrow_l \tilde{y} \quad \Rightarrow \quad \Pr_{y \leftarrow \tilde{y}}[\Phi(y_1, y)] = \Pr_{y \leftarrow \tilde{y}}[\Phi(y_2, y)]$$

For completeness, we also provide a correctness criterion for leakage cancelling.

Definition 7 (Correct pre-processing)**.** *A probabilistic program p_0 is a correct pre-processing for a deterministic program p iff for every initial memory x*

$$\langle p, x \rangle \Downarrow_l y \wedge \langle p_0; p, x \rangle \Downarrow_{l'} \tilde{y} \Longrightarrow \tilde{y} = \mathbf{1}_y$$

where $\mathbf{1}_y$ is the Dirac distribution assigning probability 1 to y and 0 to all other elements.

Finding a preprocessing program p_0 that satisfies the leakage cancelling conditions is often simple, and the practical benefits have been extensively compared in [17], [26]–[28]. For example, for sorting and all algorithms based on array comparisons, leaking the results of these comparisons can be cancelled by randomly pre-shuffling the array (assuming that all elements are distinct). For this optimization, it is critical that there exist (relatively) efficient oblivious shuffling protocols for the random shuffle operator (without any leakage). In platforms such as Sharemind [13], these protocols are offered as probabilistic instruction extensions at the source level, of the otherwise deterministic source language.

Automated verification of leakage cancelling is left for future work. Deterministic leakage can be enforced by an adaptation of the information flow type system for deterministic programs and the remaining conditions

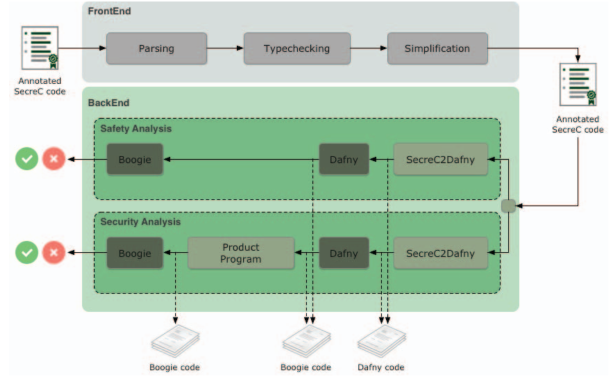

Figure 9: SecreC verification infrastructure.

can be established using probabilistic relational Hoare logic [29], using tools like EasyCrypt [30].

We conclude with a discussion of how to extend our secure compilation to encompass leakage cancelling. Lifting this result for probabilistic programs is out of the scope of this paper. However, there is a simple extension for probabilistic preprocessing scenarios. Since the preprocessing stage has no leakage, by leakage cancelling this is also the case for the final composed protocol. Assuming the existence of an atomic preprocessing protocol, the extended compiler simply prepends it to the compiled protocol. The leakage cancelling theorem then implies the existence of a cryptographic simulator that works as follows: sample any input and execute the program p to obtain some leakage; then run the simulator described in Theorem 3 to obtain a simulated trace. Intuitively, by leakage cancelling any input will lead to a leakage trace that is distributed exactly as in the real-world, and hence the simulator for the compiled protocol suffices.

IX. Implementation

This section describes our verification infrastructure for source-level analysis of deterministic MPC specifications: the *frontend* translates SecreC programs into an intermediate language close to Figure 2; the *backend* deploys our source-level analysis on top of the Dafny-Boogie verification toolchain. The architecture of our verification infrastructure is illustrated in Figure 9. The development is available at https://github.com/haslab/SecreC.

Using our infrastructure, A non-MPC expert is able to construct a SecreC program, annotate it and prove a leakage upper-bound for the program that justifies all the declassify statements as a function of the input. MPC experts typically come into play in the case of leakage cancelling, which at the moment is not supported by our verification infrastructure. For code

without declassify statements, the type system trivially suffices for guaranteeing security. Security verification is required for snippets involving declassify statements. Nevertheless, guaranteeing cryptographic security of compiled protocols programs requires proving safety and termination of SecreC programs, which is a classical deductive verification process.

A. Frontend

a) Integration: Our verification infrastructure supports the SecreC language bundled with the Sharemind SDK. This enables programmers to use our verification tools together with the Sharemind interpreter, compiler, secure execution engine and cloud deployment services. The internal operation of the Sharemind system is much more complex than the formal view of the compiler presented in this paper, as it was designed to allow for a high-degree of generality and flexility with respect to low-level protocols, data types and operations. For instance, it allows linking external protocols that are secure according to Definition 5. The main performance distinction between our formal language and Sharemind-bundled SecreC lies in the fact that we do not capture an optimization that groups operations in a SIMD style in order to save communication rounds. Even so, the general principles of the Sharemind operation match the compilation strategy we have described in our formalization, as can be seen by the cryptographic security arguments that support the system [13], which are given at a comparable (if not higher) level of abstraction.

b) Language: Much like C++, SecreC supports high-level programming features such as procedures, arrays, templates or recursion, and domain-specific support for array programming and security type polymorphism. Seeing SecreC programs as specifications of ideal secure functionalities, it becomes natural to express the security properties directly in the SecreC language. For that purpose, we have extended SecreC with an annotation language inspired by Dafny [31], a general-purpose verification language with support for procedures, loops, arrays, user-defined datatypes and native collection theories.

c) Typechecker: We have implemented a parser and a typechecker for our extended SecreC language in Haskell, and the typechecking algorithm for security type polymorphism and templates greatly resembles the treatment of ad-hoc polymorphism and type classes in Haskell. After typechecking, we apply a series of SecreC-to-SecreC simplification steps, such as removing implicit (subtyping) coercions or inlining template applications.

B. Backend

The backend translates an annotated SecreC intermediate program into two complementary Dafny programs: the first encodes functional correctness; the second assumes functional correctness and encodes security.

a) Functional embedding: As we have seen above, cryptographic MPC specifications must be (by definition) total, and our notions of source-level security are termination-insensitive. Therefore, our functional embedding of SecreC into Dafny always checks that a SecreC program is safe. It preserves the original program structure and is almost one-to-one, reducing the functional correctness of SecreC programs to that of the Dafny embedding. Under the hood, the Dafny verifier checks for functional correctness by translating to Boogie code.

b) Security embedding: Our SecreC specification language allows programmers to express leakage upper bounds and their flow through a program as annotations using the `public` keyword. These look like standard assertions but have a relational interpretation in the Security Hoare Logic from Section III.

The security analysis explores the existing translation from Dafny to Boogie to propagate security properties from SecreC to Boogie programs. We adapt the constant-time verification approach from [14] and implement a Boogie-to-Boogie transformation that computes a product program (for a SecreC program with public control flow). In the Boogie input language, procedures are defined as a sequence of basic blocks that start with a label, contain straight-line statements, and may jump at the end. For each procedure, we make shadow copies of program variables and duplicate all statements inside basic blocks to mention shadow variables instead, with two exceptions: i. procedure call statements are converted to single statements calling the product procedure with twice as many inputs and outputs; and ii. security assertions are translated to relational assertions expressing first-order logic formulas that relate original and shadowed variables, by translating `public(e)` expressions to equality expressions `e == e.shadow`.

X. Experiments

We have evaluated our infrastructure by analyzing existing SecreC specifications that are publicly available as part of the Sharemind SDK.[10] Some of these examples leak information that is subsequently cancelled using oblivious shuffling as described in Section VIII. Here, the source analyser plays an important role, as it permits

[10] https://github.com/sharemind-sdk/secrec

checking that the specification satisfies a leakage upper-bound that is compatible with the leakage cancelling theorem statement. In the remaining examples, proving a leakage bound permits matching the specification to the application requirements. We now demonstrate how to verify leakage bounds using our tool. We also discuss how the leakage cancelling steps can be performed manually, at the moment with no tool support.

a) Quick sort: Comparison-based sorting is a very heavy operation to execute obliviously in a naive way, due to the high number of oblivious branches that it involves. However any sorting specification that declassifies the results of comparing vector elements, but nothing more, gives rise to leakage that can be cancelled using oblivious shuffling of the vector prior to sorting [17]. We have proven a deterministic quick sort [26] (Section II) safe and secure with the Φ_{cmp} security policy from Section III that leaks all comparisons between array elements. For cancelling this leakage, it suffices to show that an oblivious shuffle for an array of distinct elements induces a uniform distribution on Φ_{cmp} (Theorem 4), and that the final sorted array is the same regardless of the relative ordering of the inputs (Definition 7). The algorithm can be generalized to arbitrary arrays by using the index of each element in the input list as a tie-breaker [17].

b) Gaussian elimination: Our more intricate case-study is an implementation of Gaussian elimination [28]:

```
secret uint maxFirstLoc(secret float[[1]] vec) {
    secret float best = vec[0]; secret uint idx = 0;
    for (uint i = 1; i < size(vec); i=i+1) {
        secret bool c = vec[i] > best;
        best = choose(c,vec[i],best);
        idx = choose(c,i,idx);
    }
    return declassify(idx); }
```

The algorithm receives a $k * k$ square matrix and an array of k coefficients, for $k > 0$, and solves a system of k linear equations by iterating over the columns of the matrix. For each column j, it finds the row i of the pivot (the first maximum absolute value) using a procedure that performs all comparisons obliviously, and declassifies the output. Then, it shuffles the rows i and j, performs standard matrix arithmetic on the values of the underlying rows, repeating this process until the $k-1$-th column. We first proved that the source-level declassification trace, constructed by instrumenting the source program with ghost code, consists of a permutation of the row indices. We then proved that the vector value indexed by the output of the `maxFirstLoc` function is constant for all possible permutations of the input vector. By combining these results we can easily derive that the oblivious shuffle pre-processing is sufficient to cancel the leakage under the results presented in Section VIII.

c) Radix sort: As an alternative to comparison-based sorting, we have proven the safety and security of an oblivious radix sort [26]. The implementation has an outer loop that iterates over the bit representation of the vector elements and, for each bit, operates as follows: it randomly shuffles the vector, computes a permutation that sorts the array according to the i-th bit, declassifies the permutation, and applies it in public to the vector. We prove a leakage upper-bound for the loop body that exactly matches the leakage of the permutation induced by the i-th bit. The leakage cancelling theorem then implies that the random shuffling preprocessing is sufficient to cancel this leakage in each iteration. By composition of the loop body, we get security for the entire algorithm.

d) Frequent itemsets: Finally, we have analyzed a frequent itemset algorithm that searches for co-occurring items in transactional data. Given a boolean matrix encoding of items occurring in transactions, it computes all the itemsets up to a given size k whose frequency is above a certain threshold f, revealing those itemsets. Concretely, we studied the apriori algorithm from [32], which is based on level-wise search. The algorithm computes all itemsets of size 1 up to k, using the itemsets of size $k-1$ and cached numbers of occurrences thereof to compute the itemsets of size k. For efficiency, it declassifies all the comparison tests of wether itemsets of size up to k are above the frequency. We have proven that its SecreC implementation is secure with a leakage upper bound that releases only whether every itemset up to k is frequent. Although not exercised, we could additionally prove a functional correctness property that the algorithm publishes in declassified form exactly all the frequent itemsets up to k. This match between declassified output and leakage upper bound means that the leakage of intermediate computations is benign and hence the performance benefit comes at no additional security cost.

e) Benchmarks: To give an idea of the complexity of the algorithms and the required verification effort, we have measured the number of lines of code (LOC) and the number of proof obligations (PO) for the Boogie code generated from a user-annotated SecreC program (Table I).[11] The verification experience is in all similar to deductive verification environments such as Dafny, and requires typical programmer-supplied annotations for procedure contracts and loop invariants. Since security properties often depend on auxiliary functional correctness properties, we distinguish the verification

[11]Our tool generates a single Boogie file, including encodings of standard SecreC functions and Dafny builtin theories, that are implicitly imported and replicated for all examples. Thus, we measure only proof obligations originating from each SecreC example file.

SecreC	LOC	PO		Time (s)	
		F	S	F	S
quick-sort	101	161	464	2.182	3.484
radix-sort	135	322	911	1.618	3.223
apriori	414	3877	6104	100.335	19.692
gaussian	178	610	-	2.789	-

Table I: Verification results.

effort for security annotations (S) from their required functional correctness annotations (F). We have also measured the average execution time for discharging the proof obligations over series of 10 runs on a standard MacBook Pro 2016 clocked at 2,9 GHz.

The number of security POs tends to be twice as large as the functional correctness ones, because our Boogie product program transformation duplicates all functions and statements without security annotations. For apriori, we have placed a significant effort in proving the soundness of the caching process and the observation that all frequent itemsets are reachable by adding single items to discovered frequent itemsets, as these are crucial for justifying in the security analysis that leaked information indeed corresponds to itemsets from the original database. To prove security of our gaussian example, it remains to show that our manually-instrumented leakage trace exactly matches the program's leakage trace. We could then elegantly tie source-level leakage as the permutation of the input defined by our instrumented trace. To support this, we are currently extending our implementation to handle output-dependent reasoning as in [14].

XI. Related work

Some authors use language-based verification methods for optimizing MPCs. [33] relies on epistemic modal logic to infer public intermediate values. [34] proved a similar approach sound and complete for a simple functional language. Others focus on the specification and security enforcement of MPC protocols. [5], [35], develop a domain-specific language for writing low-level SecreC protocols and a sound data flow analysis to prove the security of generated protocols w.r.t. Definition 5. [6], [36] develop the Wysteria domain-specific language to write mixed-mode MPC protocols, and give an embedding into F*. Their approach is similar to ours in that the computational behavior of programs is given by a single-threaded and a multi-threaded semantics, which are formally related. However, their notion of security is left implicit and unrelated to the simulation-based notions. We also remark that our language works at a different level of abstraction. It specifies a MPC program as a composition of high-level MPC operations (that compute an ideal functionality typically agnostic of parties). Orthogonally, the languages of [5] and [6] describe lower-level MPC protocols that give meaning to executing a MPC operation in a distributed cryptographic environment (they have notions of shares and parties). These can be integrated as primitive operations in our language. Recently, Haagh et. al [37] have provided a machine-checked proof of a concrete MPC protocol against active adversaries. They present non-interference definitions that adapt our low-level definitions to the active case, but do not consider the problem of compilation from high-level specifications.

Our approach is based on the same idea of secure compilation explored in [12]. On top of the differences highlighted in the introduction, they consider a slightly weaker security model where the adversary cannot select the initial shares. Moreover, they do not address the problem of verifying leakage of source programs. In our approach, leakage is made explicit and verified at source-level, providing early feedback to developers.

Leakage cancelling is a standard technique for optimizing MPC programs [17], [26]–[28]. However, these works are cast in the cryptographic setting, not supported by language-based verification methods. Leakage cancelling has also been considered for oblivious RAM [38] and secure hardware [39]. Both works yield provable guarantees; however, their setting is different.

XII. Conclusions

We gave a language-based security treatment of secret sharing-based MPC software stacks. We showed that our notion of source-level security is propagated, via security-aware compilation, to cryptographically secure protocols. We also provided both a formalization and an implementation of a verification technique for source-level leakage analysis in real-world examples.

Our results leave room for a number of future research directions. One interesting open problem is to investigate novel language-based approaches to deal with active adversaries. Another promising direction (following our formalization) is to implement a certified compiler or to prove an existing compiler correct. Finally, another interesting line of work is to extend our theoretical framework and tools to fully cover probabilistic specifications.

Acknowledgment

The fourth author is financed by the COMPETE 2020 Programme within project POCI-01-0145-FEDER-006961, by the FCT within project UID/EEA/50014/2013 and grant SFRH/BPD/121389/2016. The second author is financed by Project NanoSTIMA/NORTE-01-0145-FEDER-000016 through the NORTE 2020 Programme.

REFERENCES

[1] I. Damgård, K. Damgård, K. Nielsen, P. S. Nordholt, and T. Toft, "Confidential benchmarking based on multiparty computation," in *FC 2016*. Springer, 2016, pp. 169–187.

[2] L. Kamm and J. Willemson, "Secure floating point arithmetic and private satellite collision analysis," *International Journal of Information Security*, vol. 14, no. 6, pp. 531–548, 2015.

[3] M. S. Riazi, C. Weinert, O. Tkachenko, E. M. Songhori, T. Schneider, and F. Koushanfar, "Chameleon: A hybrid secure computation framework for machine learning applications," *arXiv preprint arXiv:1801.03239*, 2018.

[4] D. Bogdanov, P. Laud, and J. Randmets, "Domain-polymorphic programming of privacy-preserving applications," in *PLAS 2014*. ACM, 2014, p. 53.

[5] P. Laud and J. Randmets, "A domain-specific language for low-level secure multiparty computation protocols," in *CCS 2015*. ACM, 2015, pp. 1492–1503.

[6] A. Rastogi, M. A. Hammer, and M. Hicks, "Wysteria: A programming language for generic, mixed-mode multiparty computations," in *S&P 2015*. IEEE, 2014, pp. 655–670.

[7] C. Liu, Y. Huang, E. Shi, J. Katz, and M. Hicks, "Automating efficient ram-model secure computation," in *S&P 2014*. IEEE, 2014, pp. 623–638.

[8] A. Ben-David, N. Nisan, and B. Pinkas, "FairplayMP: a system for secure multi-party computation," in *CCS 2008*. ACM, 2008, pp. 257–266.

[9] C. Liu, X. S. Wang, K. Nayak, Y. Huang, and E. Shi, "ObliVM: A programming framework for secure computation," in *S&P 2015*. IEEE, 2015, pp. 359–376.

[10] W. Henecka, A.-R. Sadeghi, T. Schneider, and I. Wehrenberg, "TASTY: tool for automating secure two-party computations," in *CCS 2010*. ACM, 2010, pp. 451–462.

[11] A. Schropfer, F. Kerschbaum, and G. Muller, "L1-an intermediate language for mixed-protocol secure computation," in *COMPSAC 2011*. IEEE, 2011, pp. 298–307.

[12] J. C. Mitchell, R. Sharma, D. Stefan, and J. Zimmerman, "Information-flow control for programming on encrypted data," in *CSF 2012*. IEEE, 2012, pp. 45–60.

[13] D. Bogdanov, S. Laur, and J. Willemson, "Sharemind: A framework for fast privacy-preserving computations," in *ESORICS 2008*. Springer, 2008, pp. 192–206.

[14] J. B. Almeida, M. Barbosa, G. Barthe, F. Dupressoir, and M. Emmi, "Verifying constant-time implementations." in *USENIX 2016*, 2016, pp. 53–70.

[15] J. B. Almeida, M. Barbosa, G. Barthe, H. Pacheco, V. Pereira, and B. Portela, "Enforcing ideal-world leakage bounds in real-world secret sharing mpc frameworks," Cryptology ePrint Archive, Report 2018/4944, 2018, https://eprint.iacr.org/2018/4962.

[16] D. Bogdanov, P. Laud, S. Laur, and P. Pullonen, "From input private to universally composable secure multi-party computation primitives," in *CSF 2014*. IEEE, 2014, pp. 184–198.

[17] K. Hamada, R. Kikuchi, D. Ikarashi, K. Chida, and K. Takahashi, "Practically efficient multi-party sorting protocols from comparison sort algorithms," in *ICISC 2012*. Springer, 2012, pp. 202–216.

[18] A. Sabelfeld and D. Sands, "Declassification: Dimensions and principles," *Journal of Computer Security*, vol. 17, no. 5, pp. 517–548, 2009.

[19] M. Backes and B. Pfitzmann, "Computational probabilistic non-interference," in *ESORICS 2002*. Springer, 2002, pp. 1–23.

[20] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *FOCS 2001*. IEEE, 2001, pp. 136–145.

[21] I. Damgård and J. Nielsen, "Universally composable efficient multiparty computation from threshold homomorphic encryption," in *CRYPTO 2003*. Springer, 2003, pp. 247–264.

[22] V. B. Kukkala, J. S. Saini, and S. Iyengar, "Privacy preserving network analysis of distributed social networks," in *Information Systems Security*. Springer, 2016, pp. 336–355.

[23] G. Couteau, "Revisiting covert multiparty computation," *IACR Cryptology ePrint Archive*, vol. 2016, p. 951, 2016.

[24] ——, "Efficient secure comparison protocols," *IACR Cryptology ePrint Archive*, vol. 2016, p. 544, 2016.

[25] G. Barthe, T. Espitau, B. Grégoire, J. Hsu, and P.-Y. Strub, "Proving uniformity and independence by self-composition and coupling," in *LPAR 2017*. In print, 2017.

[26] D. Bogdanov, S. Laur, and R. Talviste, "A practical analysis of oblivious sorting algorithms for secure multi-party computation," in *NordSec 2014*. Springer, 2014, pp. 59–74.

[27] S. Tople, H. Dang, P. Saxena, and E.-C. Chang, "PermuteRam: Optimizing oblivious computation for efficiency," 2015. [Online]. Available: http://www.comp.nus.edu.sg/~shruti90/papers/permuteram.pdf

[28] D. Bogdanov, L. Kamm, S. Laur, and V. Sokk, "Rmind: a tool for cryptographically secure statistical analysis," *IEEE Transactions on Dependable and Secure Computing*, 2016.

[29] G. Barthe, B. Grégoire, and S. Z. Béguelin, "Formal certification of code-based cryptographic proofs," in *POPL 2009*. ACM, 2009, pp. 90–101.

[30] G. Barthe, F. Dupressoir, B. Grégoire, C. Kunz, B. Schmidt, and P. Strub, "Easycrypt: A tutorial," in *FOSAD 2012*, vol. 8604. Springer, 2013, pp. 146–166.

[31] K. R. M. Leino, "Dafny: An automatic program verifier for functional correctness," in *LPAR 2010*. Springer, 2010, pp. 348–370.

[32] D. Bogdanov, R. Jagomägis, and S. Laur, "Privacy-preserving histogram computation and frequent itemset mining with sharemind," Cybernetica research report T-4-8, Tech. Rep., 2009.

[33] F. Kerschbaum, "Automatically optimizing secure computation," in *CCS 2011*. ACM, 2011, pp. 703–714.

[34] A. Rastogi, P. Mardziel, M. Hicks, and M. A. Hammer, "Knowledge inference for optimizing secure multi-party computation," in *PLAS 2013*. ACM, 2013, pp. 3–14.

[35] M. Pettai and P. Laud, "Automatic proofs of privacy of secure multi-party computation protocols against active adversaries," in *CSF 2015*. IEEE, 2015, pp. 75–89.

[36] A. Rastogi, N. Swamy, and M. Hicks, "Wys*: A verified language extension for secure multi-party computations," *arXiv preprint arXiv:1711.06467*, 2017.

[37] H. Haagh, A. Karbyshev, S. Oechsner, B. Spitters, and P.-Y. Strub, "Computer-aided proofs for multiparty computation with active security," in *CSF 2018*. IEEE, 2018, p. In print.

[38] S. Tople, H. Dang, P. Saxena, and E.-C. Chang, "Permuteram: Optimizing oblivious computation for efficiency," Cryptology ePrint Archive, Report 2017/885, 2017, https://eprint.iacr.org/2017/885.

[39] O. Ohrimenko, M. Costa, C. Fournet, C. Gkantsidis, M. Kohlweiss, and D. Sharma, "Observing and preventing leakage in mapreduce," in *CCS 2015*. ACM, 2015, pp. 1570–1581.

APPENDIX

APPENDIX A - FULL PROOF OF THEOREM 3

To prove (ℓ, p)-privacy, we must show that, for all $\bar{x}$ and $1 \leq i \leq n$, the distributions of $(\bar{x}, y, y_i, t_i, c_i, \mathsf{st})$ are identical in the real and ideal worlds. For clarity in presentation, we present the three main definitions that are necessary for the proof in game-based form. Figure 10 presents correctness of local execution, i.e. that y_i is

uniquely defined by (x_i, t_i, c_i). Figure 11 describes low-level correctness of Π according to Definition 3, namely the correctness of Π with respect to its idealized version p. Figure 12 presents low-level security of Π according to Definition 4, namely the distributional equivalence of two runs of Π for inputs of the same leakage.

game $\mathbf{Real}_{\Pi,\mathcal{A}}()$:	**game** $\mathbf{Ideal}_{\Pi,\mathcal{A}}()$:
$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y_i, y, t_i, c_i, \mathsf{st})$	$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y'_i, y, t_i, c_i, \mathsf{st})$

Figure 10: Correctness of local execution.

game $\mathbf{Real}_{\Pi,\mathcal{A}}()$:	**game** $\mathbf{Ideal}_{p,\mathcal{A}}()$:
$(\bar{x}, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, \cdot, \cdot) \twoheadleftarrow \Pi(\bar{x})$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y, \mathsf{st})$	$(\bar{x}, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $x \leftarrow \mathrm{Unshare}(\bar{x})$ $y \leftarrow p(x)$ Return $\mathcal{A}_2(\bar{x}, y, \mathsf{st})$

Figure 11: Protocol correctness.

game $\mathsf{L}_{\Pi,\mathcal{A}}()$:	**game** $\mathsf{R}_{\Pi,\mathcal{A}}()$:

$(\bar{x}, \bar{x}', i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$
$x \leftarrow \mathrm{Unshare}(\bar{x})$
$x' \leftarrow \mathrm{Unshare}(\bar{x}')$
If $\ell(x) \neq \ell(x')$:
  Return $(b \twoheadleftarrow \{0,1\})$
$(\cdot, t, c) \twoheadleftarrow \Pi(\bar{x})$
Return $\mathcal{A}_2(\bar{x}, \bar{x}', t_i, c_i, \mathsf{st})$
	$(\bar{x}, \bar{x}', i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$
	$x \leftarrow \mathrm{Unshare}(\bar{x})$
	$x' \leftarrow \mathrm{Unshare}(\bar{x}')$
	If $\ell(x) \neq \ell(x')$:
	  Return $(b \twoheadleftarrow \{0,1\})$
	$(\cdot, t', c') \twoheadleftarrow \Pi(\bar{x}')$
	Return $\mathcal{A}_2(\bar{x}, \bar{x}', t'_i, c'_i, \mathsf{st})$

Figure 12: Probabilistic non-interference.

Proof. Our proof is a sequence of three game hops, represented in Figure 13 and described as follows.

G0 exactly matches the real world in Figure 8. G1 replaces the value of the received y_i by an alternative y'_i, calculated from (x_i, t_i, c_i) via Π_i. We upper bound the difference between these two experiments by constructing an adversary $\mathcal{B}$ against Π_i of Π such that

$$|\Pr[\mathsf{G1}() \Rightarrow \mathsf{T}] - \Pr[\mathsf{G0}() \Rightarrow \mathsf{T}]| = \mathsf{Adv}^{\Pi_i}_{\Pi,\mathcal{B}}$$

Adversary $\mathcal{B}$ executes as follows. $\mathcal{B}_1$ runs $\mathcal{A}_1$ to construct $(\bar{x}, i, \mathsf{st})$ and selects it as input for the experiment of Figure 10. This will produce a tuple $(\bar{x}, y_i, y, t_i, c_i, \mathsf{st})$, exactly matching the one that must be provided to $\mathcal{A}_2$. $\mathcal{B}_2$ returns the result obtained from $\mathcal{A}_2$.

G2 replaces the value of the received y by an alternative y', computed via p. We upper bound the difference between the two experiments by constructing an adversary $\mathcal{C}$ against p-low-correctness of Π so that

$$|\Pr[\mathsf{G2}() \Rightarrow \mathsf{T}] - \Pr[\mathsf{G1}() \Rightarrow \mathsf{T}]| = \mathsf{Adv}^{p\text{-ll-corr}}_{\Pi,p,\mathcal{C}}$$

Adversary $\mathcal{C}$ executes as follows. $\mathcal{C}_1$ runs $\mathcal{A}_1$ to construct $(\bar{x}, i, \mathsf{st})$, selecting $(\bar{x}, \mathsf{st})$ as the input for the low-level correctness experiment of Figure 11. This will produce a tuple $(\bar{x}, y, \mathsf{st})$. $\mathcal{C}_2$ will then run $(\cdot, t, c) \twoheadleftarrow \Pi(\bar{x}); y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ to produce the tuple $(\bar{x}, y'_i, y', t_i, c_i, m)$ to be given to $\mathcal{A}_2$. $\mathcal{C}_2$ returns the result obtained from $\mathcal{A}_2$.

game $\mathsf{G0}_{\Pi,\mathcal{A}}()$:	**game** $\mathsf{G1}_{\Pi,\mathcal{A}}()$:
$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y_i, y, t_i, c_i, \mathsf{st})$	$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y'_i, y, t_i, c_i, \mathsf{st})$
game $\mathsf{G1}_{\Pi,\mathcal{A}}()$:	**game** $\mathsf{G2}_{\Pi,p,\mathcal{A}}()$:
$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ $y \leftarrow \mathrm{Unshare}(\bar{y})$ Return $\mathcal{A}_2(\bar{x}, y'_i, y, t_i, c_i, \mathsf{st})$	$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ $x \leftarrow \mathrm{Unshare}(\bar{x})$ $y' \leftarrow p(x)$ Return $\mathcal{A}_2(\bar{x}, y'_i, y', t_i, c_i, \mathsf{st})$
game $\mathsf{G2}_{\Pi,p,\mathcal{A}}()$:	**game** $\mathsf{G3}_{\Pi,p,\mathcal{A}}()$:
$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $x \leftarrow \mathrm{Unshare}(\bar{x})$ $(\bar{y}, t, c) \twoheadleftarrow \Pi(\bar{x})$ $y'_i \leftarrow \Pi_i(x_i, t_i, c_i)$ $y' \leftarrow p(x)$ Return $\mathcal{A}_2(\bar{x}, y'_i, y', t_i, c_i, \mathsf{st})$	$(\bar{x}, i, \mathsf{st}) \twoheadleftarrow \mathcal{A}_1()$ $x \leftarrow \mathrm{Unshare}(\bar{x})$ $x' \twoheadleftarrow \mathsf{Find}(\ell(x))$ $\bar{x}' \twoheadleftarrow \mathsf{Share}(x')$ $(\bar{y}, t', c') \twoheadleftarrow \Pi(\bar{x}')$ $y'_i \leftarrow \Pi_i(x_i, t'_i, c'_i)$ $y' \leftarrow p(x)$ Return $\mathcal{A}_2(\bar{x}, y'_i, y', t'_i, c'_i, \mathsf{st})$

Figure 13: Proof hops for Theorem 3.

G3 replaces the values of (t, c) by those produced by an alternate protocol execution, over a set of shares whose leakage is the same as that of the original input. This makes use of $\mathsf{Find}(l)$, which we assume to be an efficient computation of a value x', whose leakage is l. We upper bound the difference between the two experiments by constructing an adversary $\mathcal{D}$ against ℓ-low-security of Π such that

$$|\Pr[\mathsf{G3}() \Rightarrow \mathsf{T}] - \Pr[\mathsf{G2}() \Rightarrow \mathsf{T}]| = \mathsf{Adv}^{\ell\text{-ll-sec}}_{\Pi,\ell,\mathcal{D}}$$

Adversary $\mathcal{D}$ executes as follows. $\mathcal{D}_1$ runs $\mathcal{A}_1$ to construct $(\bar{x}, i, \mathsf{st})$. Then, it recovers x, runs $\mathsf{Find}(\ell(x))$ to obtain an alternative x' with the same leakage, generates shares $\bar{x}'$ and selects $(\bar{x}, \bar{x}', i, x')$ as the input for the low-level security experiment of Figure 12. This will produce a tuple $(\bar{x}, \bar{x}', t_i, c_i, x')$, and $\mathcal{D}_2$ will run $y'_i \leftarrow \Pi_i(x_i, t'_i, c'_i); y' \leftarrow p(x')$ to produce the tuple $(\bar{x}, y'_i, y', t_i, c_i, \mathsf{st})$ to be given to $\mathcal{A}_2$. $\mathcal{D}_2$ returns the result obtained from $\mathcal{A}_2$.

To conclude, we have that

$$\begin{aligned}\mathbf{Adv}_{\Pi,\mathcal{F},\mathcal{A},\mathcal{S}}() &= \mathsf{Adv}^{\Pi_i}_{\Pi,\mathcal{B}} + \mathsf{Adv}^{p\text{-ll-corr}}_{\Pi,p,\mathcal{C}} + \mathsf{Adv}^{\ell\text{-ll-sec}}_{\Pi,\ell,\mathcal{D}} \\ &= 0\end{aligned}$$

and Theorem 3 follows. □