Garbling Schemes

Mihir Bellare[1] Viet Tung Hoang[2] Phillip Rogaway[2]

[1] Dept. of Computer Science and Engineering, University of California, San Diego, USA
[2] Dept. of Computer Science, University of California, Davis, USA

May 14, 2012

Abstract. Garbled circuits, a classical idea rooted in the work of Andrew Yao, have long been understood as a cryptographic *technique*, not a cryptographic *goal*. Here we cull out a primitive corresponding to this technique. We call it a *garbling scheme*. We provide a provable-security treatment for garbling schemes, endowing them with a versatile syntax and multiple security definitions. The most basic of these, *privacy*, suffices for two-party secure function evaluation (SFE) and private function evaluation (PFE). Starting from a PRF, we provide an efficient garbling scheme achieving privacy and we analyze its concrete security. We next consider *obliviousness* and *authenticity*, properties needed for private and verifiable outsourcing of computation. We extend our scheme to achieve these ends. We provide highly efficient blockcipher-based instantiations of both schemes. Our treatment of garbling schemes presages more efficient garbling, more rigorous analyses, and more modularly designed higher-level protocols.

Keywords: Garbled circuits, garbling schemes, provable security, secure function evaluation, Yao's protocol.

Table of Contents

1 Introduction

Overview. This paper is about elevating garbled circuits from a cryptographic *technique* to a cryptographic *goal*. While circuit garbling has traditionally been viewed as a method for achieving SFE (secure function evaluation) or some other cryptographic goal, we view it as an end goal in its own right, defining *garbling schemes* and formalizing several notions of security for them, these encompassing *privacy*, *authenticity*, and *obliviousness*. This enables more modular use of garbled circuits in higher-level protocols and grounds follow-on work, including the development of new and highly efficient schemes.

History. The idea of a garbled circuit is due to A. Yao, who described the technique in oral presentations [17, p. 27] about SFE [52, 53]. The first written account of the method is by Goldreich, Micali, and Wigderson [18]. The protocol they describe, crediting Yao [52], involves associating two *tokens* to each wire of a boolean circuit, these having hidden semantics of 0 and 1. Means are then provided to propagate tokens across a gate, preserving the hidden semantics. More specifically, there's a four-row table for each gate of the circuit, each row employing public-key encryption[3] to encrypt a pair of random strings whose xor is the token for the outgoing wire.

The term *garbled circuit*[4] is from Beaver, Micali, and Rogaway [10], where the method was first based on a symmetric primitive. Garbled circuits took on a modern, PRF-based instantiation in work by Naor, Pinkas, and Sumner on privacy-preserving auctions [41].

Yao's idea has been enormously impactful, engendering numerous applications, implementations, and refinements.[5] Still, there has been little definitional attention paid to garbled circuits themselves. A 2004/2009 paper by Lindell and Pinkas [34, 36] provides the first proof of Yao's protocol—to the extent one can say that a particular scheme is Yao's—but, even there, the authors do not formalize garbled circuits or what it means to securely create one. Instead, they prove that a particular garbled-circuit-using protocol, one based on double encryption,[6] is a secure two-party SFE. Implemented SFE methods do not coincide with what's in Lindell and Pinkas [36], and absence of a good abstraction boundary makes daunting the task of providing a full proof for what's actually in optimized SFE implementations.

Scattered throughout the enormous literature dealing with garbled circuits, several papers do work to abstract out what these provide. A first set of such work begins with Feige, Kilian, and Naor [14] and is followed by [8, 12, 27, 29]. Each paper aims to modularly use garbled circuits in some intending application. To that end, they single out, definitionally, precisely what they need, usually arriving at something close to what we will later call "prv.sim security over Φ_{circ}." None of the papers pick up definitions from the other, nor prove that any particular construction satisfies the notion given. A second line of definitions begins with Ishai and Kushilevitz [24] and continues with [2, 3, 5, 6, 25, 26, 47]. These works define various flavors of *randomized encodings*. Their authors do see randomized encodings as a general-purpose primitive, and the definitions elegantly support a variety of theory-centered work. However, they lack the fine-grained syntax that we shall need to investigate obliviousness, authenticity, and precise measures of efficiency. Finally, in concurrent work, Kamara and Wei offer definitions to support their idea of garbling *structured* circuits [28]. See Appendix A for further discussion of selected related work.

Contributions. We formalize what we call a *garbling scheme*. The notion is designed to support a burgeoning and practical area: the myriad applications of garbled circuits. Our definitions and result enable easy and widespread applications with modular, simplified, and yet more rigorous proofs of security.

[3] It seems to have been almost forgotten that garbled circuits were originally conceived as a technique based on public-key techniques. Abadi and Feigenbaum (1990), for example, explain that an advantage of their approach is that only one composite $N = pq$ is needed for the entire circuit, not a different one for each gate [1]. Garbled circuits have long since lost their association to public-key encryption, let alone a specific public-key technique.

[4] Synonyms in the literature include *encrypted circuit* and *scrambled circuit*.

[5] There are more than 2,700 Google-scholar-known citations to [52, 53].

[6] This approach for making the rows of the garbled gate is first mentioned by Goldreich [17].

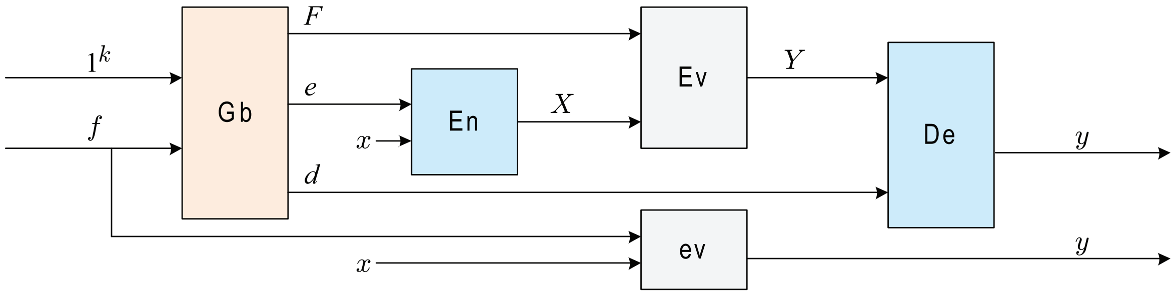

Fig. 1. Components of a garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$. Function Gb maps f and k to (F, e, d), strings encoding the garbled function, the encoding function, and the decoding function. Possession of e and x lets one compute the garbled input $X = \mathsf{En}(e, x)$; having F and X lets one calculate the garbled output $Y = \mathsf{Ev}(F, X)$; and knowing d and Y lets one recover the final output $y = \mathsf{De}(d, Y)$, which must equal $\mathsf{ev}(f, x)$.

Roughly said, a garbling algorithm Gb is a randomized algorithm that transforms a function $f : \{0,1\}^n \to \{0,1\}^m$ into a triple of functions $(F, e, d) \leftarrow \mathsf{Gb}(f)$. We require that $f = d \circ F \circ e$. The *encoding function* e turns an *initial input* $x \in \{0,1\}^n$ into a *garbled input* $X = e(x)$. Evaluating the *garbled function* F on the garbled input X gives a *garbled output* $Y = F(X)$. The *decoding function* d turns the garbled output Y into the *final output* $y = d(Y)$, which must coincide with $f(x)$. Informally, one has probabilistically factored f into $d \circ F \circ e$. Formally, it is problematic to regard Gb as operating on functions. Thus a garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ is regarded as a five-tuple of algorithms, with strings d, e, f, and F interpreted as functions under the auspices of functions De, En, ev, and Ev. See Fig. 1.

Our syntactic framework is representation-independent; circuits are nowhere to be found. One can garble DFAs, RAMs, TMs, whatever; definitionally, this isn't even seen.

Of course none of this says anything about the desired security notion. We define several. The most important is *privacy*: a party acquiring (F, X, d) shouldn't learn anything impermissible beyond that which is revealed by knowing just the final output y. To formalize "that which it is permissible to reveal" a *side-information function*, Φ, parameterizes the definition; an adversary should be able to ascertain from (F, X, d) nothing beyond $\Phi(f)$ and y. By varying Φ one can encompass the customary setting for SFE (let $\Phi(f) = f$; circuit f is not concealed) and PFE (let $\Phi(f)$ be the number of gates of f; leak just the circuit's size). We formalize privacy in multiple ways, giving an indistinguishability definition, prv.ind, and a simulation-based one, prv.sim. We show that whether or not they are equivalent depends on the side-information function Φ. For the most important ones the notions are equivalent (in general, they are not).

We provide and prove correct a simple garbling scheme, Garble1, for achieving privacy. The protocol is conveniently described in terms of a *dual-key cipher*. We show how to instantiate this primitive using a pseudorandom function (PRF), and how then to realize this PRF using a conventional blockcipher, say AES128. In this way we obtain a provably secure, blockcipher-based garbling scheme where circuit evaluation takes two AES calls per gate. The scheme is described with uncustomary precision, including a detailed and precise definition of circuits, which we deem important to ease practical implementations and gain efficiency.

We go on to suggest a still more efficient instantiation for the dual-key cipher, one where evaluating a garbled circuit needs only *one* AES128 call per gate and all blockcipher invocations use the *same* key. This is the fastest approach now known for garbling circuits. We do not prove or implement such a scheme secure in the current paper; see the discussion below.

Beyond privacy we consider *obliviousness*: a party acquiring F and X, but not d, shouldn't learn anything about f, x, or y. As with privacy, we formalize obliviousness in different but "usually" equivalent

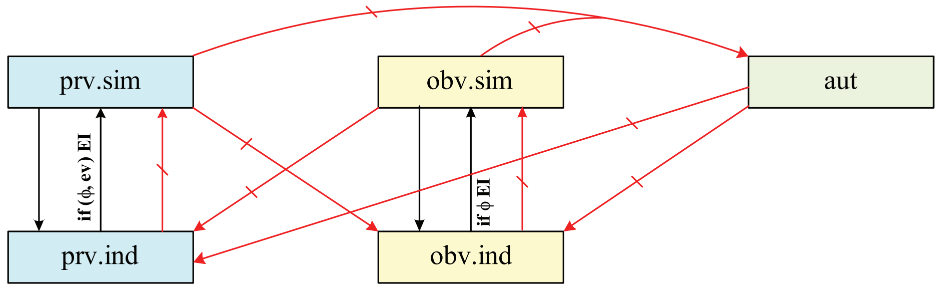

Fig. 2. Relations among security notions. A solid arrow is an implication; an if-labeled arrow, a conditional implication; a hatched arrow, a separation. Implications are found in are in Section 4; separations are in Appendix 4.

ways. Next we explore *authenticity*: a party who learns F and X should be unable to produce a garbled output Y^* different from $Y = F(X)$ that is deemed to be valid: $d(Y^*) \neq \bot$. Our interest in obliviousness and authenticity was sparked by Gennaro, Gentry, and Parno [16]; the notions arise in the context of private, verifiable outsourcing of computation.

We prove implications and separation among all security notions we have mentioned, painting a complete picture of definitions for this space. See Fig. 2.

We define a protocol, Garble2, to simultaneously achieve privacy, obliviousness, and authenticity. The assumption required is the same as before. The scheme is only a bit more complex than Garble1, the efficiency, only a little worse.

Discussion. Once viewed as a “theoretical” approach for multiparty computation, a long line of work, beginning with Fairplay [38], has made clear that circuit garbling is now a practical technique. State-of-the-art implementations by Huang *et al.* and Kreuter *et al.* can handle complex functionalities and hundreds of millions of gates [22, 23, 32]. We aim to support such work, and applications further afield. With a protocol’s garbling scheme delineated, implementations can more reasonably offer proofs for the actual scheme employed, the “messy” optimizations stripped of surrounding interaction and protocol aims. In general, an approach where the garbling scheme is conceptually separated from its use seems essential for managing complexity in this domain. As an analog, authenticated encryption took off after it was reconceptualized as a primitive, not a method formed of encryption schemes and MACs.

Garble1 and Garble2 are close to numerous other protocols (especially [41]) that incarnate Yao’s idea. Given this, one might assume that, once good definitions *are* written down, proving security would be easy, based on prior work [36]. From our experience, this is not the case; the proofs we provide are not implicit in prior work.

One thing novel about our schemes is that they admit efficient AES-based instantiations whose quantitative security may be inferred via the concrete security bounds associated to our theorems. In the past, SFE schemes supported by proofs would use objects less efficiently realizable in practice [36], or, for practical realizations, would abandon proven-secure schemes and use hash-based ones, sometimes with an unproven claim that security is maintained in the random-oracle model. Given the increasing ubiquity of AES hardware support, we believe that optimized, proven, blockcipher-based schemes are a good direction.

This paper is the first of several we envision. In it we aim to instill fresh, practice-oriented foundations in an area where, historically, omitted definitions and proofs have been the norm. The current work maintains a circumscribed focus: to investigate the definitions central to the reconceptualization of garbling schemes as a *sui generis* cryptographic goal. Upcoming work will explore several directions:

- We can construct *extremely efficient* garbling schemes, like the one-call, fixed-key, AES128-based scheme we mentioned. This can be done in a way that does not preclude the free-xor and row-

Protocol	Application	Needs	Over	Also needs
Y86 [17]	2-party SFE (semi-honest)	prv	Φ_{circ}	oblivious transfer (OT)
AF90 [1]	PFE (semi-honest)	prv	Φ_{size}	OT
FKN94 [14]	server-aided SFE (semi-honest)	prv	Φ_{circ}	none
NPS99 [41]	privacy-preserving auctions	prv	Φ_{circ}	proxy OT
KO04 [29]	2-party SFE (malicious)	prv	Φ_{circ}	OT, ZK proofs, commitment, trapdoor perm
FAZ05 [15]	private credit checking	prv	Φ_{size}	sym encryption
FM06 [39]	2-party SFE (malicious)	prv	Φ_{circ}	OT, commitment
AL07 [7]	2-party SFE (covert)	prv	Φ_{circ}	OT, commitment
LP07 [35]	2-party SFE (malicious)	prv	Φ_{circ}	OT, commitment
GKR08 [19]	one-time programs	prv!!	Φ_{size}	none (model provides one-time memory)
GMS08 [20]	2-party SFE (covert)	prv	Φ_{circ}	OT, commitment, PRG, CR hash
BFK+09 [9]	private medical diagnostics	obv	Φ_{circ}	OT, sym encryption, homomorphic enc
PSS09 [42]	private credit checking	prv	Φ_{topo}	sym encryption
BHHI10 [8]	KDM encryption	prv	Φ_{size}	KDM encryption (wrt to linear functions)
GGP10 [16]	secure outsourcing	aut! + obv!	Φ_{circ}	fully homomorphic encryption (FHE)
HS10 [21]	2-party guaranteed SFE	prv	Φ_{circ}	OT, auth encryption, asym enc, signature
SS10 [47]	worry-free encryption	prv	Φ_{size}	asym encryption, signature
Ap11 [2]	KDM encryption	prv	Φ_{size}	KDM encryption (wrt to projections)
KMR11 [27]	server-aided SFE (malicious)	aut + obv	Φ_{circ}	coin-tossing protocol, commitment
LP11 [37]	2-party SFE (malicious)	prv	Φ_{circ}	OT, commitment

Fig. 3. Recasting protocols in more generic terms. All of the above protocols appear to be alternatively describable from a garbling scheme meeting our definitions. All but [16] need the scheme to be projective.

elimination techniques that have proven so effective [22, 30, 45]. Proofs remain complex, even in the random-permutation model. Implementations are underway, these achieving about 15 nsec/gate.

- We can generalize security to the *adaptive* (=*dynamic*) setting. This is needed for one-time programs [19] and secure outsourcing [16]. For one flavor of adaptivity, prv!/obv!/aut!, the input x may depend on the garbled function F. For finer-grained notions, prv!!/obv!!/aut!!, each bit of x can depend on previously acquired X_i-values. Transformations turn prv/obv/aut schemes into prv!/obv!/aut! ones and these into prv!!/obv!!/aut!! ones.
- Building on the oft-described metaphor of lockboxes and keys (eg, [36, pp. 163–164]), we can formulate garbling-scheme security using a formal treatment of dual-key enciphering. We choose to do this by absorbing the functionality of the ideal primitive into the code-based definition. Privacy, obliviousness, and authenticity become yes/no matters—no probabilities.[7]

For all of these directions, the framework developed here serves as the needed starting point.

A thesis underlying our definitions is that they *work*—that most (though not all) applications described as using garbled circuits can be built from an arbitrary garbling scheme, instead. To date we have surveyed 20 papers containing protocols that can be recast to use a generic garbling scheme. See Fig. 3. In all cases we gain in simplicity and modularity. Applications benefit from the increased efficiency of our garbling schemes. The improvement is particularly marked in the application to KDM encryption (security with respect to key-dependent messages), where use of our abstraction leads to substantial efficiency gains over the use of the abstractions in previous work [2, 8].

[7] In fact, the only hint of an intended model for Yao's work on two-party SFE is an idealized one supporting perfect, deterministic, public-key encryption [53, Section 3.2]. Formal treatments may be possible beyond garbling schemes, to the applications that routinely use them.

2 Preliminaries

This section provides basic notation, definitions and conventions. A reader might skip this on first reading and refer back as necessary.

2.1 Notation

We let $\mathbb{N}$ be the set of positive integers. A *string* is a finite sequence of bits and $\perp$ is a formal symbol that is not a string. If A is a finite set than $y \twoheadleftarrow A$ denotes selecting an element of A uniformly at random and assigning it to y. If A is an algorithm then $A(x_1, \ldots; r)$ denotes the output of A on inputs $x_1, \ldots$ and coins r, while $y \leftarrow A(x_1, \ldots)$ means we pick r uniformly at random and let $y \leftarrow A(x_1, \ldots; r)$. We let $[A(x_1, \ldots)]$ denote the set of y that have positive probability of being output by $A(x_1, \ldots)$. We write $\mathrm{Func}(a, b)$ for $\{f \colon \{0,1\}^a \to \{0,1\}^b\}$. Polynomial time (PT) is always measured in the length of *all* inputs, not just the first. (But random coins, when singled out as an argument to an algorithm, are never regarded as an input.) As usual, a function $\varepsilon : \mathbb{N} \to \mathbb{R}^+$ is *negligible* if for every $c > 0$ there is a K such that $\epsilon(k) < k^{-c}$ for all $k > K$.

2.2 Code-based games

Our definitions and proofs are expressed via code-based games [11] so we recall here the language and specify the particular conventions we use. A code-based game—see Fig. 5 for an example—consists of an Initialize procedure, procedures that respond to adversary oracle queries, and a Finalize procedure. All procedures are optional. In an execution of game G with an adversary $\mathcal{A}$, the latter is given input 1^k where k is the security parameter, and the security parameter k used in the game is presumed to be the same. Procedure Initialize, if present, executes first, and its output is input to the adversary, who may now invoke other procedures. Each time it makes a query, the corresponding game procedure executes, and what it returns, if anything, is the response to $\mathcal{A}$'s query. The adversary's output is the input to Finalize, and the output of the latter, denoted $G^{\mathcal{A}}(k)$, is called the output of the game. Finalize may be absent in which case it is understood to be the identity function, so that the output of the game is the output of the adversary. We let "$\mathrm{Gm}^{\mathcal{A}}(k) \Rightarrow c$" denote the event that this game output takes value c and let "$\mathrm{Gm}^{\mathcal{A}}$" be shorthand for "$\mathrm{Gm}^{\mathcal{A}}(k) \Rightarrow \mathsf{true}$." Boolean flags are assumed initialized to false and $\mathsf{BAD}(\mathrm{Gm}^{A})$ is the event that the execution of game Gm with adversary A sets flag *bad* to true.

2.3 Circuits

While our definitions are representation-independent, the garbling schemes we will specify assume a circuit-based representation: the string f given to Gb gets interpreted, under ev, as encoding a circuit. Here we specify our conventions and definitions for circuits.

This are several reasons why it is important to precisely and elegantly define circuits (which, for many reasons, are not just DAGs). First, there are many "boundary cases" where only conventions can decide if something is or is not a valid circuit.[8] The boundary cases matter; we have repeatedly found that degenerate or under-analyzed circuit types materially impact if a garbling scheme is correct.[9] Beyond

[8] For example: Can an input wire be an output wire? Can an output wire be an incoming wire to another gate? Can an output wire be used twice in forming the output? Can a wire twice feed a gate? Can constants feed a gate? Can gates compute asymmetric functions like $G(x, y) = \overline{x} \vee y$?

[9] For example, the scheme of Naor, Pinkas, and Sumner [41] cannot handle a wire being used twice as an input to another gate (as when making a NOT gate from a NAND), a restriction that is nowhere explicitly said. The scheme of Beaver, Micali, and Rogaway [10] was buggy [49] because of a dependency in gate-labels associated to fan-out ≥ 2 gates.

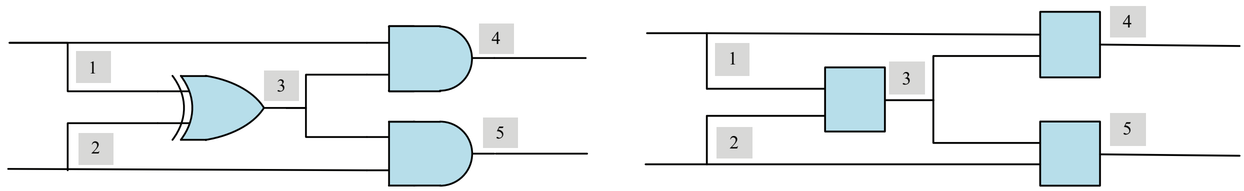

Fig. 4. Left: A conventional circuit $f = (n, m, q, A, B, G)$. It has $n=2$ inputs, $m=2$ outputs, and $q=3$ gates. Gates are numbered 3, 4, 5, according to their outgoing wires. The diagram encodes $A(3)=1$, $B(3)=2$, $A(4)=1$, $B(4)=3$, $A(5)=3$, and $B(5)=2$. The gate symbols indicate that $G_1(\cdot,\cdot) = \text{XOR}$ and $G_2(\cdot,\cdot) = G_3(\cdot,\cdot) = \text{AND}$. **Right: A topological circuit** f^- corresponding to the circuit on the left.

this, a lack of agreement on what a circuit *is* makes even informal discourse problematic.[10] Finally, we have found that it is simply not possible to properly specify a circuit-garbling algorithm Gb or a circuit-evaluation function ev, nor to carry out code-based game-playing proofs, without circuits being formalized. As an added payoff, if one establishes good conventions for circuits, then these same conventions can be used when defining a garbled circuit F and its evaluation function Ev, which we shall do.

Syntax. A (conventional) *circuit* is a 6-tuple $f = (n, m, q, A, B, G)$. Here $n \geq 2$ is the number of *inputs*, $m \geq 1$ is the number of *outputs* and $q \geq 1$ is the number of *gates*. We let $r = n + q$ be the number of *wires*. We let $\text{Inputs} = \{1, \ldots, n\}$, $\text{Outputs} = \{1, \ldots, m\}$, $\text{Wires} = \{1, \ldots, n+q\}$, $\text{OutputWires} = \{n+q-m+1, \ldots, n+q\}$, and $\text{Gates} = \{n+1, \ldots, n+q\}$. Then $A\colon\ \text{Gates} \to \text{Wires}\backslash\text{OutputWires}$ is a function to identify each gate's *first* incoming wire and $B\colon\ \text{Gates} \to \text{Wires}\backslash\text{OutputWires}$ is a function to identify each gate's *second* incoming wire. Finally $G\colon\ \text{Gates} \times \{0,1\}^2 \to \{0,1\}$ is a function that determines the *functionality* of each gate. We require $A(g) < B(g) < g$ for all $g \in \text{Gates}$. See the left side of Fig. 4 for an illustration of a circuit.

The conventions above embody all of the following. Gates have two inputs, arbitrary functionality, and arbitrary fan-out. The wires are numbered 1 to $n+q$. Every non-input wire is the outgoing wire of some gate. The ith bit of input is presented along wire i. The ith bit of output is collected off wire $n+q-m+i$. The outgoing wire of each gate serves as the name of that gate. Output wires may not be input wires and may not be incoming wires to gates. No output wire may be twice used in the output. Requiring $A_g < B_g < g$ ensures that the directed graph corresponding to f is acyclic, and that no wire twice feeds a gate; the numbering of gates comprise a topological sort.

We will routinely ignore the distinction between a circuit $f = (n, m, q, A, B, G)$ as a 6-tuple and the encoding of such a 6-tuple as a string; formally, one assumes a fixed and reasonable encoding, one where $|f|$ is $O(r \log r)$ for $r = n + q$.

Evaluating a circuit. We define a canonical evaluation function $\mathsf{ev}_{\mathsf{circ}}$. It takes a string f and a string $x = x_1 x_2 \cdots x_n$:

01 **proc** $\mathsf{ev}_{\mathsf{circ}}(f, x)$
02 $(n, m, q, A, B, G) \leftarrow f$
03 **for** $g \leftarrow n+1$ **to** $n+q$ **do** $a \leftarrow A(g)$, $b \leftarrow B(g)$, $x_g \leftarrow G(x_a, x_b)$
04 **return** $x_{n+q-m+1} \ldots x_{n+q}$

At line 03, x_a and x_b will always be well defined because of $A(g) < B(g) < g$. Circuit evaluation takes linear time. At line 02 we adopt the convention that any string f can be parsed as a circuit. (If f does

[10] For example, is there a single wire emanating from each gate, that one wire connected to all gates it feeds, or is there a separate wire from the output of a gate to each gate it feeds? (For us, it'll be the first.) These are very different meanings of *wire*.

not encode a circuit, we view it as some fixed, default circuit.) This ensures that $\mathsf{ev_{circ}}$ is well-defined for all string inputs f.

TOPOLOGICAL CIRCUITS. We say f^- is a *topological circuit* if $f^- = (n, m, q, A, B)$ for some circuit $f = (n, m, q, A, B, G)$. Thus a topological circuit is like a conventional circuit except the functionality of the gates is unspecified. See the right side of Fig. 4. Let Topo be the function that expunges the final component of its circuit-valued argument, so $f^- = \text{Topo}(f)$ is the topological circuit underlying conventional circuit f.

3 Garbling Schemes and Their Security

We define garbling schemes and security notions for them. See Section 2 should any notation seem non-obvious.

3.1 Syntax

A *garbling scheme* is a five-tuple of algorithms $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$. The first of these is probabilistic; the remaining algorithms are deterministic. A string f, the *original function*, describes the function $\mathsf{ev}(f, \cdot) : \{0,1\}^n \to \{0,1\}^m$ that we want to garble.[11] The values $n = f.n$ and $m = f.m$ depend on f and must be easily computable from it. Specifically, fix linear-time algorithms n and m to extract $f.n = \mathsf{n}(f)$ and $f.m = \mathsf{m}(f)$. [12] On input f and a security parameter $k \in \mathbb{N}$, algorithm Gb returns a triple of strings $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$. String e describes an *encoding function*, $\mathsf{En}(e, \cdot)$, that maps an *initial input* $x \in \{0,1\}^n$ to a *garbled input* $X = \mathsf{En}(e, x)$.[13] String F describes a *garbled function*, $\mathsf{Ev}(F, \cdot)$, that maps each garbled input X to a *garbled output* $Y = \mathsf{Ev}(F, X)$. String d describes a *decoding function*, $\mathsf{De}(d, \cdot)$, that maps a garbled output Y to a *final output* $y = \mathsf{De}(d, Y)$.

We levy some simple requirements on garbling schemes. First, $|F|$, $|e|$, and $|d|$ may depend only on k, $f.n$, $f.m$, and $|f|$. Formally, if $f.n = f'.n$, $f.m = f'.m$, $|f| = |f'|$, $(F, e, d) \in [\mathsf{Gb}(1^k, f)]$, and $(F', e', d') \in [\mathsf{Gb}(1^k, f')]$, then $|F| = |F'|$, $|e| = |e'|$, and $|d| = |d'|$. This is the *length* condition. Second, e and d may depend only on k, $f.n$, $f.m$, $|f|$ and the random coins r of Gb. Formally, if $f.n = f'.n$, $f.m = f'.m$, $|f| = |f'|$, $(F, e, d) = \mathsf{Gb}(1^k, f; r)$, and $(F', e', d') = \mathsf{Gb}(1^k, f'; r)$, then $e = e'$ and $d = d'$. This is the *nondegeneracy* condition. Finally, if $f \in \{0,1\}^*$, $k \in \mathbb{N}$, $x \in \{0,1\}^{f.n}$, and $(F, e, d) \in [\mathsf{Gb}(1^k, f)]$, then $\mathsf{De}(d, \mathsf{Ev}(F, \mathsf{En}(e, x))) = \mathsf{ev}(f, x)$. This is the *correctness* condition.

We say that a garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ is a *circuit garbling scheme* if $\mathsf{ev} = \mathsf{ev_{circ}}$ is the canonical circuit-evaluation function.

3.2 Projective schemes

A common approach in existing garbling schemes is for e to encode a list of *tokens*, one pair for each bit in $x \in \{0,1\}^n$. Encoding function $\mathsf{En}(e, \cdot)$ then uses the bits of $x = x_1 \cdots x_n$ to select from $e = (X_1^0, X_1^1, \ldots, X_n^0, X_n^1)$ the subvector $X = (X_1^{x_1}, \ldots, X_n^{x_n})$. Formally, we say that garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ is *projective* if for all f, $x, x' \in \{0,1\}^{f.n}$, $k \in \mathbb{N}$, and $i \in [1..n]$, when $(F, e, d) \in [\mathsf{Gb}(1^k, f)]$, $X = \mathsf{En}(e, x)$ and $X' = \mathsf{En}(e, x')$, then $X = (X_1, \ldots, X_n)$ and $X' = (X'_1, \ldots, X'_n)$ are n vectors, $|X_i| = |X'_i|$, and $X_i = X'_i$ if x and x' have the same ith bit.

---

[11] By way of example, the string f may encode a *circuit* that $\mathsf{ev}(f, \cdot)$ can evaluate at input x. Circuits and their canonical evaluation function are formalized in Section 2.

[12] For concreteness, one can define $\mathsf{n}(f)$ and $\mathsf{m}(f)$ to be n and m if f is a tuple $(n, m, \ldots)$ and define $\mathsf{n}(f) = \mathsf{m}(f) = 1$ otherwise. Of course other encoding conventions are also fine.

[13] By way of example, the encoding function e might be a sequence of $2n$ strings, called *tokens*, a pair for each bit of x. The garbled input X might then be a sequence of n strings, or *tokens*, one for each bit of x.

proc GARBLE(f_0, f_1, x_0, x_1) Game $\mathrm{PrvInd}_{\mathcal{G},\Phi}$ **if** $\Phi(f_0) \neq \Phi(f_1)$ **then return** $\perp$ **if** $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ **then return** $\perp$ **if** $\mathsf{ev}(f_0, x_0) \neq \mathsf{ev}(f_1, x_1)$ **then return** $\perp$ $b \twoheadleftarrow \{0,1\}$; $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f_b)$; $X \leftarrow \mathsf{En}(e, x_b)$ **return** (F, X, d)	**proc** GARBLE(f, x) Game $\mathrm{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ $b \twoheadleftarrow \{0,1\}$ **if** $x \notin \{0,1\}^{f.n}$ **then return** $\perp$ **if** $b = 1$ **then** $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$; $X \leftarrow \mathsf{En}(e, x)$ **else** $y \leftarrow \mathsf{ev}(f, x)$; $(F, X, d) \leftarrow \mathcal{S}(1^k, y, \Phi(f))$ **return** (F, X, d)
proc GARBLE(f_0, f_1, x_0, x_1) Game $\mathrm{ObvInd}_{\mathcal{G},\Phi}$ **if** $\Phi(f_0) \neq \Phi(f_1)$ **then return** $\perp$ **if** $\{x_0, x_1\} \not\subseteq \{0,1\}^{f_0.n}$ **then return** $\perp$ $b \twoheadleftarrow \{0,1\}$; $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f_b)$; $X \leftarrow \mathsf{En}(e, x_b)$ **return** (F, X)	**proc** GARBLE(f, x) Game $\mathrm{ObvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ $b \twoheadleftarrow \{0,1\}$ **if** $x \notin \{0,1\}^{f.n}$ **then return** $\perp$ **if** $b = 1$ **then** $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$; $X \leftarrow \mathsf{En}(e, x)$ **else** $(F, X) \leftarrow \mathcal{S}(1^k, \Phi(f))$ **return** (F, X)
proc GARBLE(f, x) $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$; $X \leftarrow \mathsf{En}(e, x)$ **return** (F, X)	**proc** FINALIZE(Y) Game $\mathrm{Aut}_{\mathcal{G}}$ **return** $(\mathsf{De}(d, Y) \neq \perp$ **and** $Y \neq \mathsf{Ev}(F, X))$

Fig. 5. Games for defining the prv.ind, prv.sim, obv.ind, obv.sim, and aut security of a garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$. Here $\mathcal{S}$ is a simulator, Φ is an information function and k is the security parameter input to the adversary. Procedure FINALIZE(b') of the first four games returns $(b = b')$.

Our definitions of security do not require schemes be projective. However, this property is needed for some important applications. For example, SFE combines a projective garbling scheme and a scheme for oblivious transfer.

3.3 Side-information functions

Privacy is rarely absolute; semantically secure encryption, for example, is allowed to reveal the length of the plaintext. Similarly, a garbled circuit might reveal the size of the circuit that was garbled, its topology (that is, the graph of how gates are connected up), or even the original circuit itself. The information that we expect to be revealed is captured by a *side-information function*, Φ, which deterministically maps f to a string $\phi = \Phi(f)$. We will parameterize our advantage notions by Φ, and in this way simultaneously define garbling schemes that may reveal a circuit's size, topology, identity, or more. We require that $f.n$ and $f.m$ be easily determined from $\phi = \Phi(f)$; formally, there must exist linear-time algorithms n' and m' that compute $f.n = \mathsf{n}'(\phi) = \mathsf{n}(f)$ and $f.m = \mathsf{m}'(\phi) = \mathsf{m}(f)$ when $\phi = \Phi(f)$.

We now specify specific side-information functions that we use in our study of circuit garbling schemes. Side-information function Φ_{size} maps a circuit $f = (n, m, q, A, B, G)$ to (n, m, q). Side-information function Φ_{topo} returns more: given f, it returns $f^- = \mathrm{Topo}(f)$. Side-information function Φ_{circ} returns most: $\Phi_{\text{circ}}(f) = f$.

3.4 Privacy

Let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a garbling scheme, $k \in \mathbb{N}$ a security parameter, and Φ a side-information function. We define an indistinguishability-based notion of privacy via game $\mathrm{PrvInd}_{\mathcal{G},\Phi}$ (top-left of Fig. 5) and a simulation-based notion of privacy via game $\mathrm{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ (top-right of Fig. 5, where $\mathcal{S}$ is a simulator). Executing either game with an adversary requires one to specify the garbling scheme, adversary, security parameter, and side-information function. Executing game PrvSim additionally requires one to specify the algorithm $\mathcal{S}$. Notation and conventions for games are specified in Section 2.

Refer first to game $\mathrm{PrvInd}_{\mathcal{G},\Phi}$. Adversary $\mathcal{A}$ gets input 1^k and must make exactly one GARBLE query. That query is answered as specified in the game, the security parameter being used here being the same as the one provided to the adversary. The adversary must eventually halt, outputting a bit b', and the game's FINALIZE procedure determines if the adversary has won on this run, namely, if $b = b'$. The corresponding advantage is defined via

$$\mathbf{Adv}_{\mathcal{G}}^{\mathrm{prv.ind},\,\Phi}(\mathcal{A},k) = 2\Pr[\mathrm{PrvInd}_{\mathcal{G},\Phi}^{\mathcal{A}}(k)] - 1,$$

the probability, normalized to $[0,1]$, that the adversary correctly predicts b. Protocol $\mathcal{G}$ is prv.ind secure over Φ if for every PT adversary $\mathcal{A}$ the function $\mathbf{Adv}_{\mathcal{G}}^{\mathrm{prv.ind},\,\Phi}(\mathcal{A},\cdot)$ is negligible.

Explaining the definition, the adversary chooses (f_0,x_0) and (f_1,x_1) such that $\Phi(f_0)=\Phi(f_1)$ and, also, $\mathsf{ev}(f_0,x_0)=\mathsf{ev}(f_1,x_1)$. The game picks challenge bit b and garbles f_b to (F,e,d). It encodes x_b as the garbled input $X=\mathsf{En}_e(x_b)$. The adversary is handed (F,X,d), which determines $y=\mathsf{De}(d,\mathsf{Ev}(F,\mathsf{En}(e,x_b)))=\mathsf{ev}(f_b,x_b)$. The adversary must guess b. In a scheme we deem secure, it should be unable to ascertain which of (f_0,x_0), (f_1,x_1) got garbled.

Next we define prv.sim security via game $\mathrm{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ associated to garbling scheme $\mathcal{G}$, information function Φ and an algorithm $\mathcal{S}$ called a simulator. The adversary $\mathcal{B}$ is run on input 1^k and must make exactly one GARBLE query. The query is answered as specified in Fig. 5, with k being the same as the input to the adversary. The adversary must eventually output a bit, and the game's FINALIZE procedure indicates if the adversary has won—again, if the adversary correctly predicted b. The adversary's advantage is

$$\mathbf{Adv}_{\mathcal{G}}^{\mathrm{prv.sim},\,\Phi,\mathcal{S}}(\mathcal{B},k) = 2\Pr[\mathrm{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k)] - 1\ ,$$

the probability, normalized to $[0,1]$, that the adversary wins. Protocol $\mathcal{G}$ is prv.sim secure over Φ if for every PT adversary $\mathcal{B}$ there is a PT algorithm $\mathcal{S}$ such that $\mathbf{Adv}_{\mathcal{G}}^{\mathrm{prv.sim},\,\Phi,\mathcal{S}}(\mathcal{B},k)$ is negligible.

Let us again explain. For the prv.sim notion we let the adversary choose (f,x). Either we garble it to $(F,e,d)\leftarrow\mathsf{Gb}(1^k,f)$ and $X\leftarrow\mathsf{En}(e,x)$, handing the adversary (F,X,d), or else we ask the simulator to devise a "fake" (F,X,d) based solely on k, $\phi=\Phi(f)$, and $y=\mathsf{ev}(f,x)$. From this limited information the simulator must produce an (F,X,d) indistinguishable, to the adversary, from the ones produced using the actual garbling scheme.

The indistinguishability definition for garbling schemes is simpler due to the absence of the simulator, but we consider this notion "wrong" when the side-information function is such that indistinguishability is inequivalent to the simulation-based definition. See Section 4.

3.5 Obliviousness

Informally, a garbling scheme achieves *obliviousness* if possession of a garbled function F and garbled input X lets one compute the garbled output Y, yet (F,X) leaks nothing about f or x beyond $\Phi(f)$. The adversary does not get the decoding function d and will not learn the output $\mathsf{De}(d,\mathsf{Ev}(F,X))$. Contrasting this with privacy, there the agent evaluating the garbled function *does* learn the output; here, she learns not even that, as a needed piece of information, d, is withheld. Privacy and obliviousness are both secrecy notions, and cut from the same cloth. Yet they will prove incomparable: a private scheme could divulge the output even without d; an oblivious scheme could reveal too much once d is shown.

As with privacy, we formalize two notions, obv.ind and obv.sim, via the games of Fig. 5. The formalizations consider games $\mathrm{ObvInd}_{\mathcal{G},\Phi}$ and $\mathrm{ObvSim}_{\mathcal{G},\Phi,\mathcal{S}}$, run with adversaries $\mathcal{A}$ and $\mathcal{B}$, respectively. As usual the adversary gets input 1^k and the security parameter used in the game is also k. The adversary makes a single call to the game's GARBLE procedure and outputs a bit b'. We define

$$\mathbf{Adv}_{\mathcal{G}}^{\mathrm{obv.ind},\,\Phi}(\mathcal{A},k) = 2\Pr[\mathrm{ObvInd}_{\mathcal{G},\Phi}^{\mathcal{A}}(k))] - 1 \quad \text{and}$$
$$\mathbf{Adv}_{\mathcal{G}}^{\mathrm{obv.sim},\,\Phi,\,\mathcal{S}}(\mathcal{B},k) = 2\Pr[\mathrm{ObvSim}_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k)] - 1$$

as the probability, normalized to $[0,1]$, that adversary's output is a correct guess of the underlying bit b. Protocol $\mathcal{G}$ is obv.ind secure over Φ if for every PT adversary $\mathcal{A}$, we have that $\mathbf{Adv}_{\mathcal{G}}^{\mathrm{obv.ind},\,\Phi}(\mathcal{A},k)$ is negligible. It is obv.sim secure over Φ if for every PT adversary $\mathcal{B}$ there exists a PT simulator $\mathcal{S}$ such that $\mathbf{Adv}_{\mathcal{G}}^{\text{obv.sim},\Phi,\mathcal{S}}(\mathcal{B},\cdot)$ is negligible.

Let us explain the difference between prv.ind and obv.ind. First, we no longer demand that $\mathsf{ev}(f, x_0) = \mathsf{ev}(f, x_1)$: the adversary may now name any (f_0, x_0) and (f_1, x_1) as long as the functions have the same side information. Second, the decoding function d is no longer provided to the adversary. The adversary must guess if (F, X) stems from garbling (f_0, x_0) or (f_1, x_1).

Similarly, the difference between prv.sim and obv.sim is two-fold. First, in the obliviousness notion the simulator is denied $y = \mathsf{ev}(f, x)$; it must create a convincing (F, X) without that. Second, the simulator no longer returns to the adversary the (simulated) decoding function d; the return value is (F, X) and not (F, X, d).

3.6 Authenticity

So far we have dealt exclusively with secrecy notions. One can formalize an authenticity property as well [16], which we do via game $\text{Aut}_{\mathcal{G}}$ of Fig. 5. Authenticity captures an adversary's inability to create from a garbled function F and its garbled input X a garbled output $Y \neq F(X)$ that will be deemed authentic.

Fix a garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$, adversary $\mathcal{A}$, and security parameter $k \in \mathbb{N}$. Run adversary $\mathcal{A}$ on input 1^k, allowing it a single call to the Garble procedure of the game. The adversary outputs a string Y, and, when it does, the game's Finalize procedure is called to decide if the adversary has won. The adversary's aut-advantage is defined as $\mathbf{Adv}_{\mathcal{G}}^{\text{aut}}(\mathcal{A}, k) = \Pr[\text{Aut}_{\mathcal{G}}^{\mathcal{A}}(k)]$. Protocol $\mathcal{G}$ is aut-secure if for all PT adversaries $\mathcal{A}$, $\mathbf{Adv}_{\mathcal{G}}^{\text{aut}}(\mathcal{A}, \cdot)$ is negligible.

3.7 Sets of garbling schemes

To compactly and precisely express relations between notions we will write them as containments and non-containments between sets of garbling schemes. To this end, for xxx $\in$ {prv.ind, prv.sim, obv.ind, obv.sim} we let $\mathsf{GS}(\text{xxx}, \Phi)$ be the set of all garbling schemes that are xxx-secure over Φ. Similarly, we let $\mathsf{GS}(\text{aut})$ be the set of all garbling schemes that are aut-secure.

We also let $\mathsf{GS}(\mathsf{ev})$ be the set of all garbling schemes $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ whose evaluation function is ev. This captures garbling schemes for a particular class of functions. As per our previous notation, $\mathsf{GS}(\mathsf{ev}_{\mathsf{circ}})$ now denotes the set of all circuit-garbling schemes.

3.8 Remarks

We end this section with discussion of our definitions.

Universal circuits. Fix one of our privacy or obliviousness notions and a projective garbling scheme $\mathcal{G}$ secure for Φ_{topo}. Then, using universal circuits, it is easy to construct a garbling scheme $\mathcal{G}_{\$}$ secure in the same sense but now with respect to Φ_{size}. This has long been understood in folklore, and is easily formalized using the language we have introduced. See Appendix B for details on universal circuits and the overhead they entail. Because of the ready translation from security with respect to Φ_{topo} to security with respect to Φ_{size}, the schemes we present in the remainder of this paper, Garble1 and Garble2, will have side-information Φ_{topo}.

Non-degeneracy. In garbling f by Gb we intend to partition f into e, F, d where e describes how to obscure the input x and where d describes how to unobscure the answer Y. We do not want $\mathsf{En}(e, \cdot)$ or $\mathsf{De}(d, \cdot)$ to actually compute $f(x)$. But this could happen if we permitted decompositions like $e = f$, $F = d = \varepsilon$, $\mathsf{En}(e, x) = \mathsf{ev}(f, x)$, and $\mathsf{Ev}(F, X) = \mathsf{De}(d, X) = X$. The nondegeneracy condition outlaws this, formalizing a sense in which e and d are independent of f. Note that we do allow e and d to depend on m, n, and even $|f|$.

Strict correctness. Our correctness condition is *strict*: you always get $\mathsf{ev}(f, x)$ by computing $\mathsf{De}(d, \mathsf{Ev}(F, \mathsf{En}(e, x)))$. One can certainly relax this requirement, and you would have to in order to regard what goes on within Lindell and Pinkas [36], say, as a garbling scheme. Yet strict correctness is not hard to achieve. Our definition could certainly be extended to say that a scheme is correct if $\Pr[(F, e, d) \leftarrow \mathsf{Gb}(1^k, f) \colon\ \mathsf{De}(d, \mathsf{Ev}(F, \mathsf{En}(e, x))) \neq \mathsf{ev}(f, x)]$ is negligible as a function of k for all f.

An undesirable way to do asymptotics. It is important not to conflate the security parameter k and f's input length n. These are conceptually distinct, and it makes perfect perfect sense to think of f, and therefore n, as fixed, while the security parameter varies. In our treatment, the security parameter k is provided to the adversary and it selects the functions to use in its attack and so, as a result, the input length n is polynomially bounded if the adversary is. The security parameter limits the input length—the input length does not define the security parameter.

Indistinguisability without side-information. The side-information function Φ does more than allow one to capture that which may be revealed by F; our prv.ind definition would be meaningless if we had effectively dialed-in $\Phi(f) = f$, the "traditional" understanding for 2-party SFE. Suppose here that we wish only to garble SHA-256, so $\mathsf{ev}(f, x) = \text{SHA-256}(x)$ for all f, x. Then the adversary can't find any distinct x_0 and x_1 such that $\mathsf{ev}(f, x_0) = \mathsf{ev}(f, x_1)$—which means that good prv.ind security will be achieved no matter *what* the garbling scheme does. An interpretation of this observation is that prv.ind is an unacceptable definition when $\Phi(f) = f$—one must ensure that less leaks about f before the definition starts to say something useful. When the adversary needs only to find $(f_0, x_0) \neq (f_1, x_1)$ such that $\mathsf{ev}(f_0, x_0) = \mathsf{ev}(f_1, x_1)$, and when Φ is designed to make sure this is an easy job for her, the definition is more meaningful.[14]

Idealized models. As in many cryptographic domains, it seems possible to obtain better efficiency working in idealized models. All of our security definitions easily lift to ideal-model settings. In the *random-oracle* model we provide any adversary or simulator, and any algorithms among the first four components of $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$, with a uniform random function $H \colon\ \{0,1\}^* \to \{0,1\}^k$. In the *ideal-cipher* model we provide, instead, an ideal cipher $E \colon\ \{0,1\}^k \times \{0,1\}^b \to \{0,1\}^b$ and its inverse $D \colon\ \{0,1\}^k \times \{0,1\}^b \to \{0,1\}^b$, each key K naming an independent random permutation $E(K, \cdot)$. In the *ideal-permutation* model we provide, instead, a random permutation $\pi \colon\ \{0,1\}^b \to \{0,1\}^b$ and its inverse $\pi^{-1} \colon\ \times\{0,1\}^b \to \{0,1\}^b$. Security results for any of these models would bound the adversary's advantage in terms of the number and type of its oracle queries.

4 Relations

We show that prv.sim always implies prv.ind, and prv.ind implies prv.sim under certain added conditions on the side-information function. We show that the same holds for obv.ind and obv.sim, under a weaker assumption on the side-information function. The conditions on the side-information function are relatively mild. We will also justify the non-implications for the security notions compactly summarized in Fig. 2. As part of this we will show that prv.ind does not always imply prv.sim and obv.ind does not always imply obv.sim.

[14] Of course this counterexample just given is concrete and our framework has been asymptotic. An asymptotic version of the counterexample is in Proposition 8.

4.1 Invertibility of side-information functions

Let Φ be a side-information function. An algorithm M is called a Φ-inverter if on input ϕ in the range of Φ it returns a preimage under Φ of that point, meaning a string f such that $\Phi(f) = \phi$. Such an inverter always exists, but it might not be efficient. We say that Φ is *efficiently invertible* if there is a polynomial-time Φ-inverter. Similarly, an algorithm M is called a (Φ, ev)-inverter if on input (ϕ, y), where $\phi = \Phi(f')$ and $y = \mathsf{ev}(f', x')$ for some f' and $x \in \{0,1\}^{f.n}$, returns an (f, x) satisfying $\Phi(f) = \phi$ and $\mathsf{ev}(f, x) = y$. We say that (Φ, ev) is *efficiently invertible* if there is a polynomial-time (Φ, ev)-inverter.

The following summarizes the invertibility attributes of the circuit-related size-information functions we defined earlier. It shows that for ev the canonical circuit evaluation, $(\Phi_{\text{size}}, \mathsf{ev})$ and $(\Phi_{\text{topo}}, \mathsf{ev})$ are efficiently invertible.

Proposition 1 *For $\Phi \in \{\Phi_{\text{size}}, \Phi_{\text{topo}}, \Phi_{\text{circ}}\}$, there is a linear-time inverter. For $\Phi \in \{\Phi_{\text{size}}, \Phi_{\text{topo}}\}$ and ev the canonical circuit-evaluation function, there is a linear-time (Φ, ev)-inverter.*

Proof (Proposition 1). We first specify a linear-time $(\Phi_{\text{topo}}, \mathsf{ev})$-inverter M_{topo}. It gets input a topological circuit f^- and an m-bit binary string $y = y_1 \dots y_m$ and proceeds as follows:

proc $M_{\text{topo}}(f^-, y)$
$(n, m, q, A, B) \leftarrow f^-$
for $(g, i, j) \in \{n+1, \dots, n+q\} \times \{0,1\} \times \{0,1\}$ **do**
  if $g \le n + q - m$ **then** $G_g(i, j) \leftarrow 0$ **else** $G_g(i, j) \leftarrow y_{g-(n+q-m)}$
$f \leftarrow (n, m, q, A, B, G)$, $x \leftarrow 0^n$
return (f, x)

We have $\text{Topo}(f) = f^-$ and $\mathsf{ev}(f, x) = y$ as desired. Next we specify a linear-time $(\Phi_{\text{size}}, \mathsf{ev})$-inverter M_{size}. It gets input (n, m, q) and an m-bit binary string $y = y_1 \dots y_m$ and proceeds as follows:

proc $M_{\text{size}}((n, m, q), y)$
for $g \in \{n+1, \dots, n+q\}$ **do** $A_g \leftarrow 1$, $B_g \leftarrow 2$
$f^- \leftarrow (n, m, q, A, B)$, $(f, x) \leftarrow M_{\text{topo}}(f^-, y)$
return (f, x)

We have $\Phi_{\text{size}}(f) = (n, m, q)$ and $\mathsf{ev}(f, x) = y$ as desired. Now a linear-time Φ_{topo}-inverter, on input $f^- = (n, m, q, A, B)$, can let $y \leftarrow 0^m$ and return $M_{\text{topo}}(f^-, y)$. Similarly, a linear-time Φ_{size}-inverter, on input (n, m, q), can let $y \leftarrow 0^m$ and return $M_{\text{size}}((n, m, q), y)$. Finally, a linear-time Φ_{circ}-inverter is trivial, returning f on input f. $\square$

In contrast, there is no efficient $(\Phi_{\text{circ}}, \mathsf{ev})$-inverter (under a computational assumption) (for ev canonical circuit evaluation); consider the case where f is drawn from a family implementing a one-way function.

4.2 Equivalence of prv.ind and prv.sim

The following says that prv.sim implies prv.ind security, and conversely if (Φ, ev) is efficiently invertible.

Proposition 2 [prv.ind $\approx$ prv.sim] For any PT Φ: (1) $\mathsf{GS}(\text{prv.sim}, \Phi) \subseteq \mathsf{GS}(\text{prv.ind}, \Phi)$ and (2) If (Φ, ev) is efficiently invertible then $\mathsf{GS}(\text{prv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \subseteq \mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$.

The first part says that if garbling scheme $\mathcal{G}$ is prv.sim secure over Φ then $\mathcal{G}$ is prv.ind secure over Φ. The second part says that if garbling scheme $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ is prv.ind secure over Φ and (Φ, ev)

is efficiently invertible then $\mathcal{G}$ is prv.sim secure over Φ. Proposition 8 proves that efficient invertibility of (Φ, ev) is required to show that prv.ind implies prv.sim, so the notions are not always equivalent.

The reductions underlying Proposition 2 are tight. This is evidenced by by Eq. (1) and Eq. (2) in the proof and the fact that the running times of the constructed adversaries or simulators are about the same as that of the starting adversary.

Proof (Proposition 2). For part (1), let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev}) \in \mathsf{GS}(\text{prv.sim}, \Phi)$. We want to show that $\mathcal{G} \in \mathsf{GS}(\text{prv.ind}, \Phi)$. Let $\mathcal{A}$ be a PT adversary attacking the prv.ind-security of $\mathcal{G}$ over Φ. We construct a PT prv.sim-adversary $\mathcal{B}$ as follows. Let $\mathcal{B}(1^k)$ run $\mathcal{A}(1^k)$. When the latter makes its query f_0, f_1, x_0, x_1 to Garble, adversary $\mathcal{B}$ returns $\perp$ to $\mathcal{A}$ if $\Phi(f_0) \neq \Phi(f_1)$ or $\mathsf{ev}(f_0, x_0) \neq \mathsf{ev}(f_1, x_1)$. Else it picks a bit c at random and queries f_c, x_c to its own Garble oracle to get back (F, X, d) and returns this to $\mathcal{A}$. The latter now returns a bit b'. Adversary $\mathcal{B}$ returns 1 if $b' = c$, $\mathsf{ev}(f_0, x_0) = \mathsf{ev}(f_1, x_1)$ and $\Phi(f_0) = \Phi(f_1)$, and 0 otherwise. Let $\mathcal{S}$ be *any* algorithm playing the role of the simulator. Then

$$\Pr\left[\,\text{PrvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid b = 1\,\right] = \frac{1}{2} + \frac{1}{2}\mathbf{Adv}^{\text{prv.ind},\Phi}_{\mathcal{G}}(\mathcal{A}, k)$$
$$\Pr\left[\,\neg\text{PrvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid b = 0\,\right] \leq \frac{1}{2}$$

where b denotes the challenge bit in game $\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$. The second claim is true because (i) if $\mathsf{ev}(f_0, x_0) = \mathsf{ev}(f_1, x_1)$ and $\Phi(f_0) = \Phi(f_1)$ then $\mathcal{S}$ has the same input regardless of c, and (ii) if $\mathsf{ev}(f_0, x_0) \neq \mathsf{ev}(f_1, x_1)$ or $\Phi(f_0) \neq \Phi(f_1)$ then $\mathcal{B}$ always answers 0, which is the correct answer in this case. Subtracting, we see that

$$\mathbf{Adv}^{\text{prv.ind},\Phi}_{\mathcal{G}}(\mathcal{A}, k) \leq 2 \cdot \mathbf{Adv}^{\text{prv.sim},\Phi,\mathcal{S}}_{\mathcal{G}}(\mathcal{B}, k)\ . \tag{1}$$

By assumption there is a PT $\mathcal{S}$ such that the RHS is negligible. Hence the LHS is negligible as well.

For part (2), let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev}) \in \mathsf{GS}(\text{prv.ind}, \Phi)$ and let M be a (Φ, ev)-inverter. We want to show that $\mathcal{G} \in \mathsf{GS}(\text{prv.sim}, \Phi)$. Let $\mathcal{B}$ be a PT adversary attacking the prv.sim-security of $\mathcal{G}$ over Φ. We define a simulator $\mathcal{S}$ that on input $1^k, y, \phi$, lets $(f, x) \leftarrow M(\phi, y)$ then $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$. It outputs $(F, \mathsf{En}(e, x), d)$. We define adversary $\mathcal{A}(1^k)$ to run $\mathcal{B}(1^k)$. When the latter makes its query f_1, x_1 to Garble, adversary $\mathcal{A}$ lets $(f_0, x_0) \leftarrow M(\Phi(f_1), \mathsf{ev}(f_1, x_1))$ and then queries f_0, f_1, x_0, x_1 to its own Garble oracle to get back (F, X, d), which it returns to $\mathcal{B}$. When the latter outputs a bit b' and halts, so does $\mathcal{A}$. Then

$$\Pr\left[\,\text{PrvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 1\,\right] = \Pr\left[\,\text{PrvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid c = 1\,\right]$$
$$\Pr\left[\,\neg\text{PrvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 0\,\right] = \Pr\left[\,\neg\text{PrvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid c = 0\,\right]$$

where b and c denote the challenge bits in games $\text{PrvInd}_{\mathcal{G},\Phi}$ and $\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$, respectively. Subtracting, we get

$$\mathbf{Adv}^{\text{prv.sim},\Phi,\mathcal{S}}_{\mathcal{G}}(\mathcal{B}, k) \leq \mathbf{Adv}^{\text{prv.ind},\Phi}_{\mathcal{G}}(\mathcal{A}, k)\ . \tag{2}$$

But the RHS is negligible by assumption, hence the LHS is as well. □

A corollary of Proposition 2 and Proposition 1 is that prv.sim and prv.ind are equivalent for circuit-garbling schemes over side-information functions Φ_{topo} and Φ_{size}, which we summarize as:

Corollary 1. $\mathsf{GS}(\text{prv.ind}, \Phi_{\text{topo}}) = \mathsf{GS}(\text{prv.sim}, \Phi_{\text{topo}})$ and $\mathsf{GS}(\text{prv.ind}, \Phi_{\text{size}}) = \mathsf{GS}(\text{prv.sim}, \Phi_{\text{size}})$.

4.3 Equivalence of obv.ind and obv.sim

The following says that obv.sim implies obv.ind security, and conversely if Φ is efficiently invertible. The invertibility condition is thus weaker than in the privacy case.

Proposition 3 [obv.ind $\approx$ obv.sim] For any PT Φ: (1) $\mathsf{GS}(\text{obv.sim}, \Phi) \subseteq \mathsf{GS}(\text{obv.ind}, \Phi)$ and (2) If Φ) is efficiently invertible then $\mathsf{GS}(\text{obv.ind}, \Phi) \subseteq \mathsf{GS}(\text{obv.sim}, \Phi)$.

Proposition 9 shows that Φ being efficiently invertible is required to show that obv.ind implies obv.sim. But the side-information function Φ we use is artificial; for any "reasonable" one we know, obv.ind and obv.sim will be equivalent.

Proof (Proposition 3). The proof is analogous to that of Proposition 2 but for completeness we provide details. For part (1), let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev}) \in \mathsf{GS}(\text{obv.sim}, \Phi)$. We want to show that $\mathcal{G} \in \mathsf{GS}(\text{obv.ind}, \Phi)$. Let $\mathcal{A}$ be a PT adversary attacking the obv.ind-security of $\mathcal{G}$ over Φ. We construct a PT obv.sim-adversary $\mathcal{B}$ as follows. Let $\mathcal{B}(1^k)$ run $\mathcal{A}(1^k)$. When the latter makes its query f_0, f_1, x_0, x_1 to Garble, adversary $\mathcal{B}$ returns $\perp$ to $\mathcal{A}$ if $\Phi(f_0) \neq \Phi(f_1)$. Else it picks a bit c at random and queries f_c, x_c to its own Garble oracle to get back (F, X, d) and returns this to $\mathcal{A}$. The latter now returns a bit b'. Adversary $\mathcal{B}$ returns 1 if $b' = c$ and $\Phi(f_0) = \Phi(f_1)$, and 0 otherwise. Let $\mathcal{S}$ be *any* algorithm playing the role of the simulator. Then

$$\Pr\left[\, \text{ObvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid b = 1 \,\right] = \frac{1}{2} + \frac{1}{2}\mathbf{Adv}^{\text{obv.ind},\,\Phi}_{\mathcal{G}}(\mathcal{A})$$
$$\Pr\left[\, \neg\text{ObvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid b = 0 \,\right] \leq \frac{1}{2}$$

where b denotes the challenge bit in game $\text{ObvSim}_{\mathcal{S}}$. The second claim is true because (i) if $\Phi(f_0) = \Phi(f_1)$ then $\mathcal{S}$ has the same input regardless of c, and (ii) if $\Phi(f_0) \neq \Phi(f_1)$ then $\mathcal{B}$ always answers 0, which is the correct answer in this case. Subtracting, we see that

$$\mathbf{Adv}^{\text{obv.ind},\,\Phi}_{\mathcal{G}}(\mathcal{A}, k) \leq 2 \cdot \mathbf{Adv}^{\text{obv.sim},\,\Phi,\mathcal{S}}_{\mathcal{G}}(\mathcal{B}, k)\ . \tag{3}$$

By assumption there is a PT $\mathcal{S}$ such that the RHS is negligible. Hence the LHS is negligible as well.

For part (2), let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev}) \in \mathsf{GS}(\text{obv.ind}, \Phi)$ and let M be a Φ-inverter. We want to show that $\mathcal{G} \in \mathsf{GS}(\text{obv.sim}, \Phi)$. Let $\mathcal{B}$ be a PT adversary attacking the obv.sim-security of $\mathcal{G}$ over Φ. we define a simulator $\mathcal{S}$ that on input $1^k, y, \phi$, lets $f \leftarrow M(\phi, y)$ then $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$. It outputs $(F, \mathsf{En}(e, x), d)$. We define adversary $\mathcal{A}(1^k)$ to run $\mathcal{B}(1^k)$. When the latter makes its query f_1, x_1 to Garble, adversary $\mathcal{A}$ lets $f_0 \leftarrow M(\Phi(f_1), \mathsf{ev}(f_1, x_1))$ and $x_0 \leftarrow 0^{f_0.n}$ and then queries f_0, f_1, x_0, x_1 to its own Garble oracle to get back (F, X, d), which it returns to $\mathcal{B}$. When the latter outputs a bit b' and halts, so does $\mathcal{A}$. Then

$$\Pr\left[\, \text{ObvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 1 \,\right] = \Pr\left[\, \text{ObvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid c = 1 \,\right]$$
$$\Pr\left[\, \neg\text{ObvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 0 \,\right] = \Pr\left[\, \neg\text{ObvSim}^{\mathcal{B}}_{\mathcal{G},\Phi,\mathcal{S}}(k) \mid c = 0 \,\right]$$

where b and c denote the challenge bits in games $\text{ObvInd}_{\mathcal{G},\Phi}$ and $\text{ObvSim}_{\mathcal{G},\Phi,\mathcal{S}}$, respectively. Subtracting, we get

$$\mathbf{Adv}^{\text{obv.sim},\,\Phi,\mathcal{S}}_{\mathcal{G}}(\mathcal{B}, k) \leq \mathbf{Adv}^{\text{obv.ind},\,\Phi}_{\mathcal{G}}(\mathcal{A}, k)\ . \tag{4}$$

But the RHS is negligible by assumption, hence the LHS is as well. ☐

Again a corollary of Proposition 3 and Proposition 1 is that obv.sim and obv.ind are equivalent for circuit-garbling schemes over side-information functions Φ_{topo} and Φ_{size}:

Corollary 2. $\mathsf{GS}(\text{obv.ind}, \Phi_{\text{topo}}) = \mathsf{GS}(\text{obv.sim}, \Phi_{\text{topo}})$ and $\mathsf{GS}(\text{obv.ind}, \Phi_{\text{size}}) = \mathsf{GS}(\text{obv.sim}, \Phi_{\text{size}})$.

4.4 Separations

We justify the non-implications for the security notions compactly summarized in Fig. 2. We state these as non-containments $\mathsf{A} \not\subseteq \mathsf{B}$ between sets of garbling schemes. We always assume $\mathsf{A} \neq \emptyset$, since otherwise the claim trivially fails.

The following says that privacy does not imply obliviousness, even when we take the strong form of privacy (simulation-style) and the weak form of obliviousness (ind-style):

Proposition 4 For all Φ and for $\mathsf{ev} = \mathsf{ev}_{\mathsf{circ}}$: $\mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{obv.ind}, \Phi)$.

Proof (Proposition 4). By assumption $\mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \neq \emptyset$ so we let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a member of this set. We construct a garbling scheme $\mathcal{G}' = (\mathsf{Gb}', \mathsf{En}, \mathsf{De}, \mathsf{Ev}', \mathsf{ev})$ such that $\mathcal{G}' \in \mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G}' \notin \mathsf{GS}(\text{obv.ind}, \Phi)$. The construction is as follows. Let $\mathsf{Gb}'(1^k, f)$ pick $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and return $((F, d), e, d)$. Let $\mathsf{Ev}'((F, d), X) = \mathsf{Ev}(F, X)$. Including d in the description of the garbled function does not harm prv.sim-security because an adversary is always given the descriptions of the garbled function and the decoding function simultaneously, so $\mathcal{G}'$ inherits the prv.sim-security of $\mathcal{G}$. On the other hand, $\mathcal{G}'$ fails to achieve obv.ind. An adversary simply makes query $(\text{OR}, \text{OR}, x_0, x_1)$ where $x_0 = 00$ and $x_1 = 11$. On receiving reply $((F, d), X)$, it outputs 0 if $\mathsf{De}(d, \mathsf{Ev}(F, X)) = \mathsf{ev}(\text{OR}, x_0)$ and 1 otherwise. This works because $0 = \mathsf{ev}(\text{OR}, x_0) \neq \mathsf{ev}(\text{OR}, x_1) = 1$ and correctness guarantees that $\mathsf{De}(d, \mathsf{Ev}(F, X)) = \mathsf{ev}(\text{OR}, x_b)$ where b is the challenge bit. $\square$

The following says that obliviousness does not imply privacy, even when we take the strong form of obliviousness (simulation-style) and the weak form of privacy (ind-style):

Proposition 5 Let $\Phi = \Phi_{\text{topo}}$ and $\mathsf{ev} = \mathsf{ev}_{\mathsf{circ}}$. Then, $\mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{prv.ind}, \Phi)$.

Proof (Proposition 5). By assumption $\mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \neq \emptyset$ so we let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a member of this set. We construct a garbling scheme $\mathcal{G}' = (\mathsf{Gb}', \mathsf{En}, \mathsf{De}', \mathsf{Ev}, \mathsf{ev})$ such that $\mathcal{G}' \in \mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G}' \notin \mathsf{GS}(\text{prv.ind}, \Phi)$. The construction is as follows. Let $\mathsf{Gb}'(1^k, f)$ pick $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and return $(F, e, (d, e))$. Let $\mathsf{De}'((d, e), Y) = \mathsf{De}(d, Y)$. Including e in the description of the decoding function does not harm obv.sim-security because an adversary is never given the description of the decoding function, so $\mathcal{G}'$ inherits the obv.sim-security of $\mathcal{G}$. On the other hand, $\mathcal{G}'$ fails to achieve prv.ind. An adversary simply makes query $(f_0, f_1, 11, 11)$ where $f_0 = \text{AND}$ and $f_1 = \text{OR}$, which is valid because $\mathsf{ev}(f_0, 11) = \mathsf{ev}(f_1, 11)$. On receiving reply $(F, X, (d, e))$, it outputs 0 if $\mathsf{De}(d, \mathsf{Ev}(F, \mathsf{En}(e, 01))) = 0$ and 1 otherwise. This works because $0 = \mathsf{ev}(f_0, 01) \neq \mathsf{ev}(f_1, 01) = 1$ and correctness guarantees that $\mathsf{De}(d, \mathsf{Ev}(F, \mathsf{En}(e, 01))) = \mathsf{ev}(f_b, 01)$ where b is the challenge bit. $\square$

The following says that privacy and obliviousness, even in conjunction and in their stronger forms (simulation-style), do not imply authenticity.

Proposition 6 For all Φ and for $\mathsf{ev} = \mathsf{ev}_{\mathsf{circ}}$: $\mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{aut})$.

Proof (Proposition 6). By assumption $\mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \neq \emptyset$ so we let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a member of this set. We construct a garbling scheme $\mathcal{G}' = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}', \mathsf{Ev}', \mathsf{ev})$ such that $\mathcal{G}' \in \mathsf{GS}(\text{prv.sim}, \Phi) \cap \mathsf{GS}(\text{obv.sim}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G}' \notin \mathsf{GS}(\text{aut})$. The construction is as follows. Let $\mathsf{Ev}'(F, X) = \mathsf{Ev}(F, X) \| 0$ and $\mathsf{De}'(d, Y \| b) = \mathsf{De}(d, Y)$ if $b = 0$ and 1 otherwise, where $b \in \{0, 1\}$. Appending a constant bit to the garbled output does not harm prv.sim security or obv.sim-security. On the other hand, $\mathcal{G}'$ fails to achieve aut. An adversary simply makes query $(\text{OR}, 00)$ and then outputs $1 \| 1$. $\square$

The following says that authenticity implies neither privacy nor obliviousness, even when the latter are in their weaker (ind style) form.

Proposition 7 Let $\Phi = \Phi_{\text{topo}}$ and $\mathsf{ev} = \mathsf{ev}_{\mathsf{circ}}$. Then $\mathsf{GS}(\text{aut}) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{prv.sim}, \Phi) \cup \mathsf{GS}(\text{obv.sim}, \Phi)$.

Proof (Proposition 7). By assumption $\mathsf{GS}(\text{aut}) \cap \mathsf{GS}(\mathsf{ev}) \neq \emptyset$ so we let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a member of this set. We construct a garbling scheme $\mathcal{G}' = (\mathsf{Gb}', \mathsf{En}, \mathsf{De}, \mathsf{Ev}', \mathsf{ev})$ such that $\mathcal{G}' \in \mathsf{GS}(\text{aut}) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G}' \notin \mathsf{GS}(\text{prv.sim}, \Phi) \cup \mathsf{GS}(\text{obv.sim}, \Phi)$. The construction is as follows. Let $\mathsf{Gb}'(1^k, f)$ pick $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and return $((F, f), e, d)$. Let $\mathsf{Ev}'((F, f), X) = \mathsf{Ev}(F, X)$. Appending f to F does not harm authenticity as the adversary has chosen f, and thus already knows it, in its attack. On the other hand, the garbled function leaks f so privacy and obliviousness both fail over Φ_{topo}. □

We saw in Proposition 2 that prv.ind implies prv.sim if (Φ, ev) is efficiently invertible. Now we show that this assumption is necessary by showing that in general prv.ind does not imply prv.sim. We say that $P\colon \{0,1\}^* \to \{0,1\}^*$ is a permutation if: (1) for every $x \in \{0,1\}^*$ we have $|P(x)| = |x|$; (2) for every distinct $x_0, x_1 \in \{0,1\}^*$ we have $P(x_0) \neq P(x_1)$. We say that P is one-way if for every PT adversary $\mathcal{I}$ the function $\mathbf{Adv}_P^{\text{ow}}(\mathcal{I}, \cdot)$ is negligible, where for each $k \in \mathbb{N}$ we have let

$$\mathbf{Adv}_P^{\text{ow}}(\mathcal{I}, k) \;=\; \Pr[\mathcal{I}(P(x)) = x] \;,$$

the probability over $x \twoheadleftarrow \{0,1\}^k$. We associate to P the evaluation function $\mathsf{ev}^P(f, x) = P(x)$ for all $f, x \in \{0,1\}^*$.

Proposition 8 Let Φ be the identity function. Let P be a one-way permutation and let $\mathsf{ev} = \mathsf{ev}^P$. Then $\mathsf{GS}(\text{prv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{prv.sim}, \Phi)$.

We note that the (Φ, ev) in Proposition 8 is not efficiently invertible due to the one-wayness of P, so this separation is consistent with Proposition 2.

Proof (Proposition 8). We build $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ so that $\mathcal{G} \in \mathsf{GS}(\text{prv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G} \notin \mathsf{GS}(\text{prv.sim}, \Phi)$. Let $\mathsf{Gb}(1^k, f) = (f, \varepsilon, \varepsilon)$ for any f. Let $\mathsf{En}(\varepsilon, x) = x$ and $\mathsf{De}(\varepsilon, Y) = Y$ for all $x, Y \in \{0,1\}^n$. We claim that $\mathbf{Adv}_{\mathcal{G}}^{\text{prv.ind}, \Phi}(\mathcal{A}) = 0$ for any (even computationally-unbounded) adversary $\mathcal{A}$. Consider an adversary $\mathcal{A}$ that makes GARBLE query (f_0, f_1, x_0, x_1). For the response to not be $\perp$ it must be that $f_0 = f_1$ and $\mathsf{ev}(f_0, x_0) = \mathsf{ev}(f_1, x_1)$, meaning $P(x_0) = P(x_1)$. Since P is a permutation, it follows that $x_0 = x_1$, and thus the advantage of the adversary must be 0. However, one can trivially break the prv.sim security of $\mathcal{G}$, with respect to any PT simulator $\mathcal{S}$ as follows. Adversary $\mathcal{A}(1^k)$ lets $f \leftarrow \varepsilon$ and $x \twoheadleftarrow \{0,1\}^k$. It then queries (f, x) to the oracle GARBLE. On receiving (F, X, d), it outputs 1 if $X = x$, and 0 otherwise. The simulator $\mathcal{S}$ gets input f and $y = \mathsf{ev}(f, x) = P(x)$ and produces (F, X, d). The probability that $X = x$ is negligible by the one-wayness of P, so the adversary's output is 1 with negligible probability when the challenge bit is 0. □

We saw in Proposition 3 that obv.ind implies obv.sim if Φ is efficiently invertible. Now we show that this assumption is necessary by showing that in general obv.ind does not imply obv.sim. Let π be a bijection from $\text{Func}(2,1)$ to $\{0,1\}^4$. Such a bijection exists, as $|\text{Func}(2,1)| = 16$. Let P be a one-way permutation. We associate to P and π the following side-information function $\Phi_{P,\pi}$. For each circuit $f = (n, m, q, A, B, G)$, let $\Phi_{P,\pi}(f) = (\text{Topo}(f), P(L))$, where $L = L_1 \ldots L_q$ and $L_i = \pi(G_{n+i})$ for each $1 \le i \neq q$.

Proposition 9 Let P be a one-way permutation and π a bijection from $\text{Func}(2,1)$ to $\{0,1\}^4$. Let $\Phi = \Phi_{P,\pi}$ and $\mathsf{ev} = \mathsf{ev}_{\mathsf{circ}}$. Then $\mathsf{GS}(\text{obv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \not\subseteq \mathsf{GS}(\text{obv.sim}, \Phi)$.

We note that the one-wayness of P means Φ is not efficiently invertible, so this separation is consistent with Proposition 3. We also note that although Φ migh look strange it is functionally equivalent to Φ_{circ} in the sense that $\Phi(f_0) = \Phi(f_1)$ iff $\Phi_{\text{circ}}(f_0) = \Phi_{\text{circ}}(f_1)$. This is true because Φ reveals the topology by definition, and since π, P are bijections, $P(\pi(G))$ uniquely determines G. This implies $\mathsf{GS}(\text{xxx}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \subseteq \mathsf{GS}(\text{xxx}, \Phi_{\text{circ}}) \cap \mathsf{GS}(\mathsf{ev})$ for both $\text{xxx} \in \{\text{obv.ind}, \text{obv.sim}\}$. (It does not imply the sets are equal because P is one-way.) On the other hand $\mathsf{GS}(\text{xxx}, \Phi_{\text{topo}}) \cap \mathsf{GS}(\mathsf{ev}) \subseteq \mathsf{GS}(\text{xxx}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ so the sets in the Proposition contain interesting and natural schemes even though they might look strange at first glance.

Proof (Proposition 9). By assumption $\mathsf{GS}(\text{obv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev}) \neq \emptyset$ so we let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a member of this set. We construct a garbling scheme $\mathcal{G}' = (\mathsf{Gb}', \mathsf{En}, \mathsf{De}, \mathsf{Ev}', \mathsf{ev})$ such that $\mathcal{G}' \in \mathsf{GS}(\text{obv.ind}, \Phi) \cap \mathsf{GS}(\mathsf{ev})$ but $\mathcal{G}' \notin \mathsf{GS}(\text{obv.sim}, \Phi)$. The construction is as follows. Let $\mathsf{Gb}'(1^k, f)$ pick $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and return $(f \| F, e, d)$. Let $\mathsf{Ev}'(f \| F, X)$ return $\mathsf{Ev}(F, X)$. We claim that $\mathcal{G}'$ is obv.ind secure over Φ but not obv.sim secure over Φ.

To justify the first claim, consider an adversary $\mathcal{A}$ that makes GARBLE query (f_0, f_1, x_0, x_1). For the response to not be $\perp$ it must be that $\Phi(f_0) = \Phi(f_1)$ and hence, by the functional equivalence noted above, that $\Phi_{\text{circ}}(f_0) = \Phi_{\text{circ}}(f_1)$. Thus $f_0 = f_1$. Prepending f to F therefore does no harm to the obv.ind security.

We justify the second claim by presenting an adversary $\mathcal{B}$ that trivially breaks the obv.sim security of $\mathcal{G}'$, with respect to any PT simulator. Adversary $\mathcal{B}(1^k)$ picks an arbitrary topological circuit f^- of k gates and chooses $L \twoheadleftarrow \{0,1\}^{4k}$. Let $L = L_1 \ldots L_k$, where each $L_i \in \{0,1\}^4$. Let $G_{n+i} = \pi^{-1}(L_i)$ for every $1 \le i \le k$, and let $f = (f^-, G)$. The adversary then queries $(f, 0^{f.n})$ to GARBLE. When receiving the reply (F', X), it returns 1 if the first $|f|$ bits of F' equal f, and returns 0 otherwise, so that it always returns 1 when the challenge bit in the game is 1. A simulator $\mathcal{S}$ gets input 1^k and $\phi = (f^-, P(L))$ and produces an output (F', X). Let $f^- = (n, m, k, A, B)$. Note that if the simulator can produce G, it also can produce $L = L_1 \ldots L_q$ with each $L_i = \pi(G_{n+i})$. The probability that the first $|f|$ bits of F' equal $f = (f^-, G)$ is therefore negligible by the one-wayness of P, because L's sampling is independent of f^-. So the adversary's output is 1 with negligible probability when the challenge bit is 0. □

5 Achieving Privacy: Garble1

We provide a simple, privacy-achieving circuit-garbling scheme. The scheme Garble1 will be described in terms of a (new) primitive that we call a dual-key cipher. We will analyze and prove correct a version of Garble1 where the dual-key cipher is built from a PRF. Instantiating the latter via AES leads to highly efficient garbling.

5.1 Dual-key ciphers

A *dual-key cipher* (DKC) $\mathbb{E}$ associates to keys $\mathtt{A}, \mathtt{B} \in \{0,1\}^k$ and tweak $\mathtt{T} \in \{0,1\}^{\tau(k)}$ a permutation $\mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}} : \{0,1\}^{\eta(k)} \to \{0,1\}^{\eta(k)}$. Let $\mathbb{D}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}} : \{0,1\}^{\eta(k)} \to \{0,1\}^{\eta(k)}$ denote the inverse of this permutation. Then it is required that the maps $\mathtt{A}, \mathtt{B}, \mathtt{T}, \mathtt{P} \mapsto \mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}(\mathtt{P})$ and $\mathtt{A}, \mathtt{B}, \mathtt{T}, \mathtt{C} \mapsto \mathbb{D}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}(\mathtt{C})$ be PT computable. We refer to τ as the tweak length and η as the block length.

The notion above is purely syntactical—we do not, at the moment, give any security definition. Our theorems will prove security for Garble1 using specific dual-key ciphers based on PRFs, the assumption being security of the underlying PRF.

5.2 Definition of Garble1

Let $\mathbb{E}$ be a dual-key cipher with tweak length τ and block length η satisfying $\eta(k) \ge k$ for all $k \in \mathbb{N}$. We associate to it the garbling scheme Garble1[$\mathbb{E}$] as shown in Fig. 6. Wires carry $\eta(k)$-bit *tokens*. A token X will encode a k-bit *key* and a 1-bit *type*. Rather arbitrarily, the key is the first k bits of the token and the type is its final bit. In the code, we write $(\mathtt{X}, \mathtt{x}) \leftarrow X$ (lines 106 and 154) to mean that $\mathtt{X}$ is the first k bits of X and $\mathtt{x}$ is the final bit, namely the lsb. When we write $\mathtt{T} \leftarrow g \,\|\, \mathtt{a} \,\|\, \mathtt{b}$, where $g \in \mathbb{N}$ and $\mathtt{a}, \mathtt{b} \in \{0,1\}$,

100 **proc** $\mathsf{Gb}(1^k, f)$
101 $(n, m, q, A, B, G) \leftarrow f$
102 **for** $i \in \{1, \ldots, n+q-m\}$ **do** $t \twoheadleftarrow \{0,1\}$, $X_i^0 \twoheadleftarrow \mathcal{X}_t$, $X_i^1 \twoheadleftarrow \mathcal{X}_{\bar{t}}$
103 **for** $i \in \{n+q-m+1, \ldots, n+q\}$ **do** $X_i^0 \twoheadleftarrow \mathcal{X}_0$, $X_i^1 \twoheadleftarrow \mathcal{X}_1$
104 **for** $(g, i, j) \in \{n+1, \ldots, n+q\} \times \{0,1\} \times \{0,1\}$ **do**
105 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
106 $\quad (\mathtt{A}, \mathtt{a}) \leftarrow X_a^i$, $(\mathtt{B}, \mathtt{b}) \leftarrow X_b^j$, $\mathtt{T} \leftarrow g \,\|\, \mathtt{a} \,\|\, \mathtt{b}$, $P[g, \mathtt{a}, \mathtt{b}] \leftarrow \mathbb{E}_{\mathtt{A},\mathtt{B}}^{\mathtt{T}}\big(X_g^{G_g(i,j)}\big)$
107 $F \leftarrow (n, m, q, A, B, P)$
108 $e \leftarrow (X_1^0, X_1^1, \ldots, X_n^0, X_n^1)$
109 $d \leftarrow \varepsilon$
110 **return** (F, e, d)

120 **proc** $\mathsf{En}(e, x)$
121 $(X_1^0, X_1^1, \ldots, X_n^0, X_n^1) \leftarrow e$
122 $X \leftarrow (X_1^{x_1}, \ldots, X_n^{x_n})$
123 **return** X

130 **proc** $\mathsf{De}(d, Y)$
131 $(Y_1, \ldots, Y_m) \leftarrow Y$
132 **for** $i \in \{1, \ldots, m\}$ **do** $y_i \leftarrow \mathrm{lsb}(Y_i)$
133 **return** $y \leftarrow y_1 \cdots y_m$

140 **proc** $\mathsf{ev}(f, x)$
141 $(n, m, q, A, B, G) \leftarrow f$
142 **for** $g \leftarrow n+1$ **to** $n+q$ **do**
143 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
144 $\quad x \leftarrow G_g(x_a, x_b)$
145 **return** $x_{n+q-m+1} \ldots x_{n+q}$

150 **proc** $\mathsf{Ev}(F, X)$
151 $(n, m, q, A, B, P) \leftarrow F$
152 **for** $g \leftarrow n+1$ **to** $n+q$ **do**
153 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
154 $\quad (\mathtt{A}, \mathtt{a}) \leftarrow X_a$, $(\mathtt{B}, \mathtt{b}) \leftarrow X_b$, $\mathtt{T} \leftarrow g \,\|\, \mathtt{a} \,\|\, \mathtt{b}$, $X_g \leftarrow \mathbb{D}_{\mathtt{A},\mathtt{B}}^{\mathtt{T}}\big(P[g, \mathtt{a}, \mathtt{b}]\big)$
155 **return** $(X_{n+q-m+1}, \ldots, X_{n+q})$

Fig. 6. Garbling scheme Garble1. Its components are $(\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ where ev, shown for completeness, is the canonical circuit evaluation. We assume a dual-key cipher $\mathbb{E}$ with tweak length τ and block length η, and let $\mathbb{D}$ denote its inverse. We write $\mathcal{X}_b$ for all length-$\eta(k)$ strings that end in a b.

we mean that $g \bmod 2^{\tau(k)-2}$ is encoded as a $(\tau(k)-2)$-bit string and $\mathtt{a} \,\|\, \mathtt{b}$ is concatenated, yielding a $\tau(k)$-bit tweak.

To garble a circuit, we begin selecting two tokens for each wire, one of each type. One of these will represent 0—the token is said to have *semantics* of 0—while the other will represent 1. The variable X_i^b names the token of wire i with semantics (not type!) of b. Thus the encoding function e (lines 120–123) will map $x = x_1 \cdots x_n \in \{0,1\}^n$ to $X = (X_1^{x_1}, \ldots, X_n^{x_n})$. For each wire i that is not an output wire, we select, at line 102, random tokens of opposite type, making the association between a token's type and its semantics random. For each wire i that is an output wire, we again select random tokens of opposite types, but this time the token's type *is* the token's semantics.

Lines 104–106 compute q garbled truth tables, one for each gate g. Table $P[g, \cdot, \cdot]$ has four rows, entry $\mathtt{a}, \mathtt{b}$ the row to use when the left incoming token is of type $\mathtt{a}$ and the right incoming token is of type $\mathtt{b}$. The token that gets encrypted for this row (end of line 106) is the token for the outgoing-wire wire with the correct semantics. At line 154, given two tokens X_a and X_b we use their types to determine which entry of the propagation table we need to decrypt. The description of the decoding function d (line 109) is empty because no information is needed to map an output token to its semantics, the type being the semantics.

In mapping a token to its key and type via $(\mathtt{X}, \mathtt{x}) \leftarrow X$ one might expect the token to have one more bit than the key, whence $\mathtt{X} \,\|\, \mathtt{x} = X$. But we have not demanded this, making it possible that the key and type fields overlap ($\mathtt{X} = X$, $\mathtt{x} = \mathrm{lsb}(\mathtt{X})$), or that a token contains bits of unspecified use.

The ev function is simply $\mathsf{ev}_{\mathsf{circ}}$, the code repeated for completeness and to make visible the common framework with Ev.

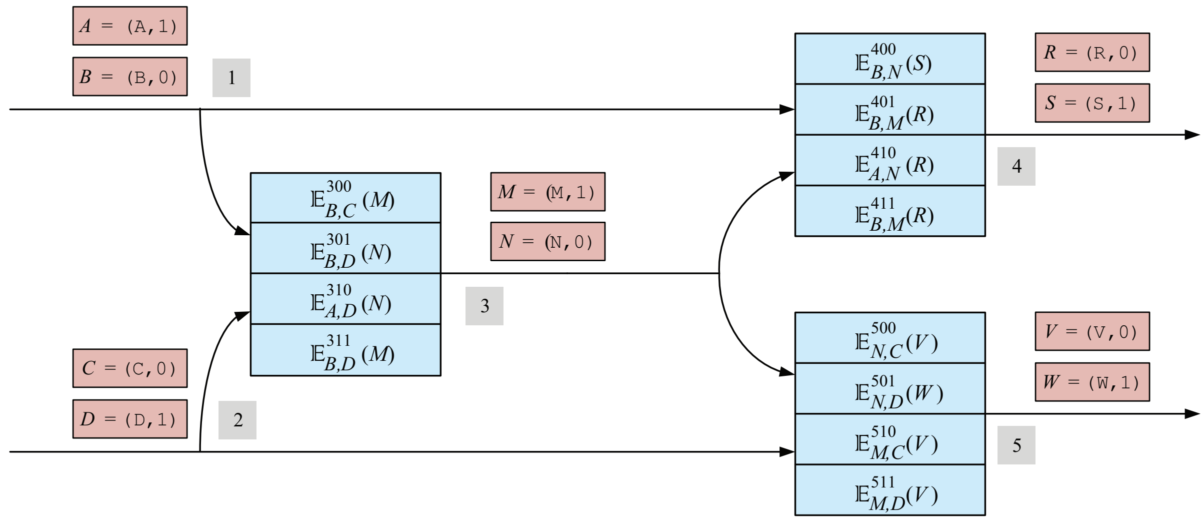

Fig. 7. Garbled circuit corresponding to the conventional circuit of Fig. 4. For each wire i, the token with semantics 0 (that is, X_i^0 is written on top; the token with semantics 1 (that is, X_i^1) is written on bottom. Possession of token A and C, for example, lets one decrypt the third row of the leftmost garbled gate (since A ends in 1 and C ends in 0) to recover token N. The final output is the concatenation of the LSBs of the output wires.

5.3 Converting a PRF to a dual-key cipher

Our primary interest will be in instantiating a dual-key cipher via a PRF. Let F associate to key $\mathtt{K} \in \{0,1\}^k$ a map $\mathsf{F}_\mathtt{K}\colon \{0,1\}^{\tau(k)} \to \{0,1\}^{k+1}$. We require that the map $\mathtt{K}, \mathtt{T} \mapsto \mathsf{F}_\mathtt{K}(\mathtt{T})$ be PT computable. We refer to τ as the input length.

The prf-advantage of an adversary $\mathcal{D}$ against F is $\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F}}(\mathcal{D}, k) = 2\Pr[\mathrm{PRF}^{\mathcal{D}}_{\mathsf{F}}(k)] - 1$ where game PRF_F is as follows. Initialize picks a random bit b and a random k-bit key $\mathtt{K}$. The adversary has access to procedure Fn that maintains a table $\mathtt{Tbl}[\cdot]$ initially everywhere undefined. Given $\mathtt{T} \in \{0,1\}^{\tau(k)}$, the procedure returns $\mathsf{F}(\mathtt{K}, \mathtt{T})$ if $b = 1$. Otherwise, it picks and returns $\mathtt{Tbl}[\mathtt{T}] \twoheadleftarrow \{0,1\}^{k+1}$ if $\mathtt{Tbl}[\mathtt{T}] = \bot$, or returns $\mathtt{Tbl}[\mathtt{T}]$ if $\mathtt{Tbl}[\mathtt{T}] \neq \bot$. Finalize$(b')$ returns $(b = b')$. We say that F is PRF-secure if $\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F}}(\mathcal{D}, \cdot)$ is negligible for all PT adversaries $\mathcal{D}$.

Given a PRF F as above, we define the dual-key cipher $\mathbb{E}$ via $\mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}(P) = \mathsf{F}_\mathtt{A}(\mathtt{T}) \oplus \mathsf{F}_\mathtt{B}(\mathtt{T}) \oplus P$. This dual-key cipher has tweak length τ and block length $\eta(k) = k + 1$ and is denoted $\mathbb{E}[\mathsf{F}]$. It is important that $k + 1 = \eta(k)$: we do not have overlap between the key and type. During evaluation, token types are revealed, so had the key contained it, it would amount to leakage of a bit of the key. In that case, PRF security would no longer suffice [43].

5.4 Security of the PRF-based scheme

We prove the privacy of Garble1$[\mathbb{E}[\mathsf{F}]]$ over Φ_{topo}. As Φ_{topo} efficiently invertible, it matters little if we understand privacy as prv.ind or prv.sim (cf. Corollary 1); the theorem below employs the former. The proof is in Section 5.7.

Theorem 10. Let F be a PRF. Then $\mathcal{G} = \mathrm{Garble1}[\mathbb{E}[\mathsf{F}]] \in \mathsf{GS}(\mathrm{prv.ind}, \Phi_{\mathrm{topo}})$. □

We emphasize that the theorem is underlain by an explicit, blackbox, uniform reduction as follows. There is a blackbox reduction U s.t. if $\mathcal{A}(1^k)$ outputs circuits of $r(k) \le 2^{\tau(k)-2}$ wires and fan-out at most $\nu(k)$ then $\mathcal{D} = U^{\mathcal{A}}$ achieves $\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F}}(\mathcal{D}, k) \ge (1/2r(k)) \cdot \mathbf{Adv}^{\mathrm{prv.ind},\,\Phi_{\mathrm{topo}}}_{\mathcal{G}}(\mathcal{A}, k)$, makes $2\nu(k)$ oracle queries, and runs in time about that of $\mathcal{A}$ plus the time for $8r(k)$ computations of F on k-bit keys. The small overhead implicit in the word "about" is manifest in the proof.

5.5 Discussion

Scheme Garble1 does not satisfy obliviousness or authenticity. To defeat obliviousness, an adversary can just make the query (AND, OR, 00, 11) to receive (F, X), and then evaluate $Y = \mathsf{Ev}(F, X)$, returning 1 if $\mathsf{De}(\varepsilon, Y) = 1$ and 0 otherwise. This adversary has advantage 1. To defeat authenticity, an adversary can query (OR, 11), and then output $(0^{k+1}, 0^{k+1})$. Again it has advantage 1. We will soon describe Garble2 that satisfies obliviousness and authenticity in addition to privacy.

The primitive used by Lindell and Pinkas [36] as a basis for encryption of gate entries is a randomized, IND-CPA secure symmetric encryption scheme additionally having properties they call *elusive range* and *efficiently verifiable range*. Dual-key ciphers, in contrast, are deterministic. Our PRF-based instantiation avoids probabilistic encryption. Besides speed it results in shorter ciphertexts for each row of each gate. The additional properties of encryption assumed by [36] are to allow the evaluator to know which gate entry is the "correct" one. Our simpler and more efficient solution via type bits (the "point-and-permute" technique) is from [41, 46].

It is possible to define security for dual-key ciphers, giving an advantage notion strong enough to guarantee security for Garble1 (and Garble2, in the next section), yet liberal enough to be provably achieved by the realizations we have mentioned, and the additional instantiations we will describe below. But such DKC security notions are not very simple, and, for now, we prefer to treat dual-key ciphers purely for their convenient syntax. This does not make the DKC notion moot; it is still the conceptually best way to understand Garble1 and Garble2.

5.6 AES-based instantiations

We now consider concrete instantiations. This means we fix a value k of the security parameter and suggest ways to realize $\mathbb{E}$ on k-bit keys based on blockciphers, specifically AES. Security for these instantiations can be derived via the concrete security bounds that we stated above following Theorem 10. Different choices of instantiation lead to different tradeoffs between assumptions and efficiency. We begin with ways to instantiate F on k-bit keys:

- Let $\mathsf{F}(\mathtt{K}, \mathtt{T})$ be the first $k+1$ bits of $E_{\mathtt{K}}(\mathtt{T} \,\|\, 0) \,\|\, E_{\mathtt{K}}(\mathtt{T} \,\|\, 1)$ for a blockcipher E having block length and key length of k; to be concrete, $E = \text{AES128}$, $k = |\mathtt{K}| = 128$, and $\tau = |\mathtt{T}| = 127$. This construction is a good PRF under the standard assumption that E is a good PRP. With this instantiation, evaluating a garbled gate costs four AES operations.
- Let $\mathsf{F}_{\mathtt{K}}(\mathtt{T})$ be $E_{\mathtt{K}\|0}(\mathtt{T})$ for a blockcipher having a $(k+1)$-bit key and block size, say $E = \text{AES128}$ and $k = |\mathtt{K}| = 127$, and $\tau = |\mathtt{T}| = 128$. Assuming that E is a good PRP is not enough to prove that F is a good PRF, as zeroing out a bit of the key does not, in general, preserve PRF security [43]. Still, it seems reasonable to directly assume this F is a good PRF. Costs are halved compared to the above; now, evaluating a garbled gate requires two AES operations.

Next we suggest some further ways to make the dual-key cipher $\mathbb{E}$ directly, meaning not via a PRF. The first follows the double-encryption realization of garbled gates attributed to Yao by Goldreich [17] (who would have been understood that primitive to be probabilistic, not a blockcipher). The second method is extremely efficient—the most efficient approach now known. Implementation work is currently underway to measure the magnitude of the gain:

- Let $\mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}(\mathtt{P}) = E_{\mathtt{A}}(E_{\mathtt{B}}(\mathtt{P}))$ (the tweak is ignored), where $E\colon \{0,1\}^\kappa \times \{0,1\}^\kappa \to \{0,1\}^\kappa$ is a blockcipher, say AES128. For a proof we would model E as an ideal cipher. Composition of encryption schemes is understood by many researchers to be Yao's original approach, although the earliest expositions make this seem doubtful.
- Let $\mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}(\mathtt{P}) = E_{\mathsf{const}}(K) \oplus K \oplus \mathtt{P}$ where $K = \mathtt{A} \oplus \mathtt{B} \oplus \mathtt{T}$ and $E = \text{AES128}$, say, and const is a fixed 128-bit string. Here $k = \tau = \eta = 128$. With this instantiation evaluating a gate costs only 1 AES

operation. Even more important, all AES operations employ a single, fixed key. This allows one to take full advantage of AES-NI hardware support to get extremely high speeds. For a proof, we would model $E_{\mathsf{const}}(\cdot)$ as a random permutation π, giving the adversary access to oracles for π and its inverse. Other one-call, fixed-key schemes are possible, for obliviousness, authenticity, and dynamic security, and adjustments to allow the free-xor and row-reduction optimizations [30, 45].

5.7 Proof of security of Garble1

Moving to the simulation-based notion, we prove the following lemma. Theorem 10 then follows via Proposition 2.

Lemma 1. Let F be a PRF. Then $\mathcal{G} = \text{Garble1}[\mathbb{E}[\mathsf{F}]] \in \mathsf{GS}(\text{prv.sim}, \Phi_{\text{topo}})$. $\square$

We will prove the following concrete statement. There are blackbox reductions U, V s.t. if $\mathcal{B}(1^k)$ outputs circuits of $q(k)$ gates, $r(k) \leq 2^{\tau(k)-2}$ wires and fan-out $\nu(k)$ then $\mathcal{D} = U^{\mathcal{B}}$ and $\mathcal{S} = V^{\mathcal{B}}$ achieve $\mathbf{Adv}_{\mathcal{G}}^{\text{prv.sim},\Phi_{\text{topo}},\mathcal{S}}(\mathcal{B}, k) \leq r(k) \cdot \mathbf{Adv}_{\mathsf{F}}^{\text{prf}}(\mathcal{D}, k)$. The running time of $\mathcal{S}$ is that of $2q(k)$ computations of F plus $O((r+q)k)$. The running time of $\mathcal{D}$ is that of $\mathcal{B}$ plus the time for $8r(k)$ computations of F on k-bit keys and $\mathcal{D}$ makes $2\nu(k)$ oracle queries.

The proof reformulates the real game in a different way and then specifies a simulator. It then presents some hybrid games and constructs the adversary $\mathcal{D}$. Let $A \| a \leftarrow X$ mean that A is the first k bits and a the last bit of $X \in \{0,1\}^{k+1}$.

Consider game Real of Fig. 8. We claim

$$\Pr\left[\,\text{Real}^{\mathcal{B}} \Rightarrow 1\,\right] \;=\; \Pr\left[\,\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k) \mid b = 1\,\right]$$

where b is the challenge bit of game $\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ of Fig. 5. To justify this, recall that in the Garble1 scheme, each wire i carries tokens X_i^0, X_i^1 with semantics 0 and 1 respectively. If wire i ends up having value (semantics) v_i in the computation $y \leftarrow \mathsf{ev}(f, x)$ then token $X_i^{v_i}$ becomes visible to the adversary while $X_i^{1-v_i}$ stays invisible. But the result is still that the adversary gets a random token, of random type on non-output wires and of type equal to the semantics on output wires. Our game aims to make this explicit. It does not pick tokens to have an a priori semantics but instead picks for each wire i a "visible" token $\mathtt{V}_i \,\|\, t_i$ and an "invisible" token $\mathtt{I}_i \,\|\, \bar{t}_i$. It then ensures that the tokens the adversary gets are the visible ones, meaning that $X_i^{v_i} = \mathtt{V}_i \,\|\, t_i$, whence the name. Thus, the game ensures that visible tokens are random except that on output wires their types correspond to the output values. At line 11 we let $\mathsf{ev}(f, x, i)$ return the bit value of wire i in the evaluation of f on input x.

The simulator of Fig. 8 gets as input $y = \mathsf{ev}(f, x)$ and $f^- = \Phi_{\text{topo}}(f)$, the topological circuit corresponding to f. It aims to produce an output similar to that of Game Real. It creates visible tokens per wire just like Game Real. It lets the output wires have the semantics given by y and lets all other wires have semantics 0. Table entries are random except for the ones that may be opened by the visible tokens.

Now consider the hybrid games Hy_ℓ of Fig. 8, defined for $0 \leq \ell \leq n + q - m$. At line 14, invisible tokens are only chosen for wires of index more than ℓ. A table entry for gate g is random if the corresponding entry in game Real uses some $\mathtt{I}_i$, with $i \leq \ell$, as an argument to F. References to invisible tokens are always well-defined; we refer to $\mathtt{I}_i$ only when $i > \ell$, due to the requirement $a < b < g$. We claim

$$\Pr\left[\,\text{Hy}_0^{\mathcal{B}} \Rightarrow 1\,\right] = \Pr\left[\,\text{Real}^{\mathcal{B}} \Rightarrow 1\,\right]$$
$$\Pr\left[\,\text{Hy}_{r-m}^{\mathcal{B}} \Rightarrow 1\,\right] = \Pr\left[\,\neg\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k) \mid b = 0\,\right]$$

where b is the challenge bit of game $\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}$ of Fig. 5 and simulator $\mathcal{S}$ in this game is that of Fig. 8. The first claim is easily verified as, when $\ell = 0$, games Hy_ℓ and Real are the same. For the second claim,

proc $\textsc{Garble}(f, x)$ — Game Real

$(n, m, q, A, B, G) \leftarrow f$
for $i \leftarrow 1$ **to** $n + q$ **do**
  $\mathtt{V}_i \twoheadleftarrow \{0,1\}^k,\ \ v_i \leftarrow \mathsf{ev}(f, x, i)$
  if $i \leq n + q - m$ **then** $t_i \twoheadleftarrow \{0,1\}$ **else** $t_i \leftarrow v_i$
  $X_i^{v_i} \leftarrow \mathtt{V}_i \,\|\, t_i,\ \ \mathtt{I}_i \twoheadleftarrow \{0,1\}^k,\ \ X_i^{\overline{v}_i} \leftarrow \mathtt{I}_i \,\|\, \overline{t}_i$
for $g \leftarrow n + 1$ **to** $n + q$ **do**
  $a \leftarrow A(g),\ \ b \leftarrow B(g)$
  $\mathtt{T} \leftarrow g t_a t_b,\ \ P[g, t_a, t_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{v_g}$
  $\mathtt{T} \leftarrow g \overline{t}_a t_b,\ \ P[g, \overline{t}_a, t_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, v_b)}$
  $\mathtt{T} \leftarrow g t_a \overline{t}_b,\ \ P[g, t_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(v_a, \overline{v}_b)}$
  $\mathtt{T} \leftarrow g \overline{t}_a \overline{t}_b,\ \ P[g, \overline{t}_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, \overline{v}_b)}$
$F \leftarrow (n, m, q, A, B, P)$
return $(F, X_1^{v_1}, \ldots, X_n^{v_n}, \varepsilon)$

algorithm $\mathcal{S}(1^k, y, f^-)$ — Simulator

$(n, m, q, A, B) \leftarrow f^-$
for $i \leftarrow 1$ **to** $n + q$ **do**
  $\mathtt{V}_i \twoheadleftarrow \{0,1\}^k$
  if $i \leq n + q - m$ **then** $t_i \twoheadleftarrow \{0,1\}$, $v_i \leftarrow 0$ **else** $t_i \leftarrow y_i$, $v_i \leftarrow y_i$
  $X_i^{v_i} \leftarrow \mathtt{V}_i \,\|\, t_i$
for $g \leftarrow n + 1$ **to** $n + q$ **do**
  $a \leftarrow A(g),\ \ b \leftarrow B(g)$
  $\mathtt{T} \leftarrow g \,\|\, t_a \,\|\, t_b,\ \ P[g, t_a, t_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{v_g}$
  $P[g, \overline{t}_a, t_b]$, $P[g, t_a, \overline{t}_b]$, $P[g, \overline{t}_a, \overline{t}_b] \twoheadleftarrow \{0,1\}^{k+1}$
$F \leftarrow (n, m, q, A, B, P)$
return $(F, X_1^{v_1}, \ldots, X_n^{v_n}, \varepsilon)$

proc $\textsc{Garble}(f, x)$ — Game Hy_ℓ

$(n, m, q, A, B, G) \leftarrow f$
for $i \leftarrow 1$ **to** $n + q$ **do**
  $\mathtt{V}_i \twoheadleftarrow \{0,1\}^k,\ \ v_i \leftarrow \mathsf{ev}(f, x, i)$
  if $i \leq n + q - m$ **then** $t_i \twoheadleftarrow \{0,1\}$ **else** $t_i \leftarrow v_i$
  $X_i^{v_i} \leftarrow \mathtt{V}_i \,\|\, t_i$
  if $i > \ell$ **then** $\mathtt{I}_i \twoheadleftarrow \{0,1\}^k,\ \ X_i^{\overline{v}_i} \leftarrow \mathtt{I}_i \,\|\, \overline{t}_i$
for $g \leftarrow n + 1$ **to** $n + q$ **do**
  $a \leftarrow A(g),\ \ b \leftarrow B(g)$
  $\mathtt{T} \leftarrow g \,\|\, t_a \,\|\, t_b,\ \ P[g, t_a, t_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{v_g}$
  if $a \leq \ell$ **then** $P[g, \overline{t}_a, t_b] \twoheadleftarrow \{0,1\}^{k+1}$ **else** $\mathtt{T} \leftarrow g \overline{t}_a t_b,\ \ P[g, \overline{t}_a, t_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, v_b)}$
  if $b \leq \ell$ **then** $P[g, t_a, \overline{t}_b] \twoheadleftarrow \{0,1\}^{k+1}$ **else** $\mathtt{T} \leftarrow g t_a \overline{t}_b,\ \ P[g, t_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(v_a, \overline{v}_b)}$
  if $a \leq \ell$ **then** $P[g, \overline{t}_a, \overline{t}_b] \twoheadleftarrow \{0,1\}^{k+1}$ **else** $\mathtt{T} \leftarrow g \overline{t}_a \overline{t}_b,\ \ P[g, \overline{t}_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, \overline{v}_b)}$
$F \leftarrow (n, m, q, A, B, P)$
return $(F, X_1^{v_1}, \ldots, X_n^{v_n}, \varepsilon)$

Fig. 8. Real game, simulator, and hybrid games (Hy_ℓ for $0 \leq \ell \leq n + q - m$) for the proof of Lemma 1.

suppose $\ell = n + q - m$. Each gate g has either $A(g) \leq n + q - m$ or $B(g) \leq n + q - m$, so all table entries are random. The visible tokens are not chosen the same way by Hy_{n+q-m} and $\mathcal{S}$, but both are uniform.

proc $\text{Garble}(f, x)$ // as defined by adversary $\mathcal{D}$
00 $(n, m, q, A, B, G) \leftarrow f,\ \ \ell \twoheadleftarrow \{1, \ldots, q+n-m\}$
10 **for** $i \leftarrow 1$ **to** $n+q$ **do**
11 $\quad \mathtt{V}_i \twoheadleftarrow \{0,1\}^k,\ \ v_i \leftarrow \mathsf{ev}(f, x, i)$
12 $\quad$ **if** $i \le n+q-m$ **then** $t_i \twoheadleftarrow \{0,1\}$ **else** $t_i \leftarrow v_i$
13 $\quad X_i^{v_i} \leftarrow \mathtt{V}_i \,\|\, t_i$
14 $\quad$ **if** $i > \ell$ **then** $\mathtt{I}_i \twoheadleftarrow \{0,1\}^k,\ \ X_i^{\overline{v}_i} \leftarrow \mathtt{I}_i \,\|\, \overline{t}_i$
20 **for** $g \leftarrow n+1$ **to** $n+q$ **do**
21 $\quad a \leftarrow A(g),\ \ b \leftarrow B(g)$
22 $\quad \mathtt{T} \leftarrow g t_a t_b,\ \ P[g, t_a, t_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{v_g}$
23 $\quad$ **if** $a < \ell$ **then** $P[g, \overline{t}_a, t_b] \twoheadleftarrow \{0,1\}^{k+1}$
24 $\quad$ **if** $a = \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, \overline{t}_a \,\|\, t_b,\ \ P[g, \overline{t}_a, t_b] \leftarrow \text{Fn}(\mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, v_b)}$
25 $\quad$ **if** $a > \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, \overline{t}_a \,\|\, t_b,\ \ P[g, \overline{t}_a, t_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{V}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, v_b)}$
26 $\quad$ **if** $b < \ell$ **then** $P[g, t_a, \overline{t}_b] \twoheadleftarrow \{0,1\}^{k+1}$
27 $\quad$ **if** $b = \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, t_a \,\|\, \overline{t}_b,\ \ P[g, t_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \text{Fn}(\mathtt{T}) \oplus X_g^{G_g(v_a, \overline{v}_b)}$
28 $\quad$ **if** $b > \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, t_a \,\|\, \overline{t}_b,\ \ P[g, t_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{V}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(v_a, \overline{v}_b)}$
29 $\quad$ **if** $a < \ell$ **then** $P[g, \overline{t}_a, \overline{t}_b] \twoheadleftarrow \{0,1\}^{k+1}$
30 $\quad$ **if** $a = \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, \overline{t}_a \,\|\, \overline{t}_b,\ \ P[g, \overline{t}_a, \overline{t}_b] \leftarrow \text{Fn}(\mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, \overline{v}_b)}$
31 $\quad$ **if** $a > \ell$ **then** $\mathtt{T} \leftarrow g \,\|\, \overline{t}_a \,\|\, \overline{t}_b,\ \ P[g, \overline{t}_a, \overline{t}_b] \leftarrow \mathsf{F}(\mathtt{I}_a, \mathtt{T}) \oplus \mathsf{F}(\mathtt{I}_b, \mathtt{T}) \oplus X_g^{G_g(\overline{v}_a, \overline{v}_b)}$
32 $F \leftarrow (n, m, q, A, B, P)$
33 **return** $(F, X_1^{v_1}, \ldots, X_n^{v_n}, \varepsilon)$

Fig. 9. Procedure Garble used by adversary $\mathcal{D}$ attacking prf-security of F, based on the adversary $\mathcal{B}$ attacking the prv.sim-security of Garble1.

Adversary $\mathcal{D}$ runs $\mathcal{B}$. When the latter makes a $\text{Garble}(f, x)$ query, it replies via the code of Fig. 9. When $\mathcal{B}$ halts with output b', adversary $\mathcal{D}$ returns b' and halts. We claim that

$$\Pr\left[\,\text{PRF}^{\mathcal{D}} \mid b = 1\,\right] = \frac{1}{n+q-m} \sum_{\ell=1}^{n+q-m} \Pr\left[\,\text{Hy}_{\ell-1}^{\mathcal{B}} \Rightarrow 1\,\right]$$

$$\Pr\left[\,\neg\text{PRF}^{\mathcal{D}} \mid b = 0\,\right] = \frac{1}{n+q-m} \sum_{\ell=1}^{n+q-m} \Pr\left[\,\text{Hy}_{\ell}^{\mathcal{B}} \Rightarrow 1\,\right]$$

where b is the challenge bit of game PRF. To justify this, first consider $b = 0$, so that Fn is a random function. The inputs $\mathtt{T}$ to which it is applied are always different because both the gate index and the types of the tokens on both wires are included in $\mathtt{T}$, so the table entries at lines 24, 27, 30 are random. On the other hand if $b = 1$ then think of Fn as playing the role of $\mathsf{F}(\mathtt{I}_\ell, \cdot)$. (It is important that $\mathcal{D}$ did not pick $\mathtt{I}_\ell$ so that this correspondence of $\mathtt{I}_\ell$ with the key underlying oracle Fn in game PRF makes sense.) Subtracting, we bound

$$\begin{aligned}
\mathbf{Adv}_{\mathsf{F}}^{\text{prf}}(\mathcal{D}) &= \Pr\left[\,\text{PRF}^{\mathcal{D}} \mid b = 1\,\right] - \Pr\left[\,\neg\text{PRF}^{\mathcal{D}} \mid b = 0\,\right] \\
&\le \frac{1}{n+q-m}\left(\Pr\left[\,\text{Hy}_0^{\mathcal{B}} \Rightarrow 1\,\right] - \Pr\left[\,\text{Hy}_{n+q-m}^{\mathcal{B}} \Rightarrow 1\,\right]\right) \\
&= \frac{\Pr\left[\,\text{Real}^{\mathcal{B}} \Rightarrow 1\,\right] - \Pr\left[\,\neg\text{PrvSim}_{\mathcal{G},\Phi,\mathcal{S}}^{\mathcal{B}}(k) \mid b = 0\,\right]}{n+q-m} \\
&= \frac{\mathbf{Adv}_{\mathcal{G}}^{\text{prv.sim},\,\Phi_{\text{topo}},\,\mathcal{S}}(\mathcal{B})}{n+q-m}\,.
\end{aligned}$$

This concludes the proof.

200 **proc** $\mathsf{Gb}(1^k, f)$
201 $(n, m, q, A, B, G) \leftarrow f$
202 **for** $i \in \{1, \ldots, n+q\}$ **do** $t \twoheadleftarrow \{0,1\}$, $X_i^0 \twoheadleftarrow \mathcal{X}_t$, $X_i^1 \twoheadleftarrow \mathcal{X}_{\overline{t}}$
203 **for** $(g, i, j) \in \{n+1, \ldots, n+q\} \times \{0,1\} \times \{0,1\}$ **do**
204 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
205 $\quad (\mathtt{A}, \mathtt{a}) \leftarrow X_a^i$, $(\mathtt{B}, \mathtt{b}) \leftarrow X_b^j$, $\mathtt{T} \leftarrow g \,\|\, \mathtt{a} \,\|\, \mathtt{b}$, $P[g, \mathtt{a}, \mathtt{b}] \leftarrow \mathbb{E}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}\big(X_g^{G_g(i,j)}\big)$
206 $F \leftarrow (n, m, q, A, B, P)$
207 $e \leftarrow (X_1^0, X_1^1, \ldots, X_n^0, X_n^1)$
208 $d \leftarrow (X_{n+q-m+1}^0, X_{n+q-m+1}^1, \ldots, X_{n+q}^0, X_{n+q}^1)$
209 **return** (F, e, d)

220 **proc** $\mathsf{En}(e, x)$
221 $(X_1^0, X_1^1, \ldots, X_n^0, X_n^1) \leftarrow e$
222 $X \leftarrow (X_1^{x_1}, \ldots, X_n^{x_n})$
223 **return** X

230 **proc** $\mathsf{De}(d, Y)$
231 $(Y_1, \ldots, Y_m) \leftarrow Y$, $(Y_1^0, Y_1^1, \ldots, Y_m^0, Y_m^1) \leftarrow d$
232 **for** $i \in \{1, \ldots, m\}$ **do**
233 $\quad$ **if** $Y_i = Y_i^0$ **then** $y_i \leftarrow 0$
234 $\quad$ **else if** $Y_i = Y_i^1$ **then** $y_i \leftarrow 1$ **else return** $\bot$
235 **return** $y \leftarrow y_1 \cdots y_m$

240 **proc** $\mathsf{ev}(f, x)$
241 $(n, m, q, A, B, G) \leftarrow f$
242 **for** $g \leftarrow n+1$ **to** $n+q$ **do**
243 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
244 $\quad x \leftarrow G_g(x_a, x_b)$
245 **return** $x_{n+q-m+1} \ldots x_{n+q}$

250 **proc** $\mathsf{Ev}(F, X)$
251 $(n, m, q, A, B, P) \leftarrow F$
252 **for** $g \leftarrow n+1$ **to** $n+q$ **do**
253 $\quad a \leftarrow A(g)$, $b \leftarrow B(g)$
254 $\quad (\mathtt{A}, \mathtt{a}) \leftarrow X_a$, $(\mathtt{B}, \mathtt{b}) \leftarrow X_b$, $\mathtt{T} \leftarrow g \,\|\, \mathtt{a} \,\|\, \mathtt{b}$, $X_g \leftarrow \mathbb{D}^{\mathtt{T}}_{\mathtt{A},\mathtt{B}}\big(P[g, \mathtt{a}, \mathtt{b}]\big)$
255 **return** $(X_{n+q-m+1}, \ldots, X_{n+q})$

Fig. 10. Garbling scheme Garble2. Its components are $(\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ where ev, shown for completeness, is the canonical circuit evaluation. We assume a dual-key cipher $\mathbb{E}$ with tweak length τ and block length η, and let $\mathbb{D}$ denote its inverse. We write $\mathcal{X}_b$ for all length-$\eta(k)$ strings that end in a b.

6 Achieving Privacy, Authenticity and Obliviousness: Garble2

We now describe a scheme Garble2 that satisfies not only privacy but also obliviousness and authenticity. The scheme is like Garble1 except, first, the last bit of a token is always uniform, even for output wires. This will give obliviousness. Next, the string encoding the decoding function is made to list all the tokens for all the output wires, ordered to make clear which tokens have what semantics. This engenders authenticity. See Fig. 10.

Talking through some of the pseudocode, line 202 now assigns a token with random semantics to each and every wire. Lines 203–207 compute the garbled function F and encoding function e exactly as with Garble1. Line 208 now records the vector of tokens for each of the m output wires. (Recall that, under our conventions, the last m of the r total wires are the output wires, these providing the m output bits, in order.) At lines 230–235 decoding procedure De, when presented a $2m$-vector d and an m-vector Y, verifies that each component of the latter is in the corresponding set of two allowed values. If so, we determine the correct semantics for this output bit using our convention that Y_i^b has semantics b.

Garble2 simultaneously achieves privacy, obliviousness, and authenticity if instantiated in the same manner as we instantiated Garble1. This is captured by the following result. Again, as per Corollary 2 it does not matter whether we consider ind or sim, and for simplicity we pick the former.

Theorem 11. Let F be a PRF. Then $\mathcal{G} = \text{Garble2}[\mathbb{E}[\mathsf{F}]] \in \mathsf{GS}(\text{prv.ind}, \Phi_{\text{topo}}) \cap \mathsf{GS}(\text{obv.ind}, \Phi_{\text{topo}}) \cap \mathsf{GS}(\text{aut})$. $\square$

As usual this asymptotic claim is underlain by concrete blackbox reductions and concrete bounds as follows. There are blackbox reductions U_{xxx} for xxx $\in \{\text{prv.ind}, \text{obv.ind}, \text{aut}\}$ s.t. if $\mathcal{A}(1^k)$ outputs circuits of $r(k) \leq 2^{\tau(k)-2}$ wires, fan-out $\nu(k)$, and achieves xxx-advantage of at least ε, then $\mathcal{D} = U_{\mathrm{xxx}}^{\mathcal{A}}$ achieves prf-advantage at least $\varepsilon/2r(k) - 2^{-k}$, makes $2\nu(k)$ oracle queries, and has running time about that of $\mathcal{A}$ plus the time for $8q(k)$ computations of F on k-bit keys.

Proof (Theorem 11). Part (a) can be proved by adapting the proof of Theorem 10 as follows. For every output wire i, in addition to the visible token $\mathtt{V}_i \parallel t_i$, the games also return the invisible token $\mathtt{I}_i \parallel \overline{t_i}$, and the type t_i is uniformly random. Lines 20-31 of the games in Fig. 8–9 might use $\mathtt{I}_i$ only if i is an incoming wire of some gate. Since no incoming wire of some gate can be an output wire, the argument in the proof of Theorem 10 still applies here.

For part (b) we again adapt the proof of Theorem 10, with the following differences. First, for every output wire i, in addition to the "visible" token $\mathtt{V}_i \parallel t_i$, the games also return the invisible token $\mathtt{I}_i \parallel \overline{t_i}$, and the type t_i is uniformly random. Next, the value v_i is always 0, and consequently, the games never need to call $\mathsf{ev}(f, x, i)$, which makes the last game in the chain of hybrid games in Fig. 8 correspond to the code of the simulator. Finally, since we are working with obliviousness, the argument of the simulator should be $1^k, f^-$, not $1^k, y, f^-$. The simulator now knows nothing about y, but since t_i and v_i are now always uniformly random, it never needs to use y.

For part (c), we construct an adversary $\mathcal{B}$ such that

$$\mathbf{Adv}^{\mathrm{aut}}_{\mathrm{Garble2}[\mathsf{F}]}(\mathcal{A}, k) \ \leq\ \mathbf{Adv}^{\mathrm{prv.ind},\,\Phi_{\mathrm{topo}}}_{\mathrm{Garble2}[\mathsf{F}]}(\mathcal{B}, k) + 2^{-k}$$

where $\mathcal{B}$'s running time is at most that of $\mathcal{A}$ plus an overhead linear to the size of $\mathcal{A}$'s query. We then apply (a) to $\mathcal{B}$. The adversary $\mathcal{B}(1^k)$ runs $\mathcal{A}(1^k)$. Suppose that $\mathcal{A}$ queries (f, x), with $f = (n, m, q, A, B, G)$. Let $y = y_1 \ldots y_m = f(x)$. Adversary $\mathcal{B}$ then constructs a circuit $f' = (n, m, q, A, B, G')$ as follows. For every gate g, if its outgoing wire j is an output wire then G'_g is a constant function that always outputs $y_{j-(n+q-m)}$. Otherwise, $G'_g = G_g$. Adversary $\mathcal{B}$ queries its oracle with (f', f, x, x). Since $f'(x) = y$ and $\Phi_{\mathrm{topo}}(f') = \Phi_{\mathrm{topo}}(f)$ and the side-information function is Φ_{topo}, the query (f', f, x, x) in game PrvInd will not result in answer $\perp$. Let (F, X, d) denote the answer. $\mathcal{B}$ gives (F, X) to $\mathcal{A}$ as response to its query (f, x). $\mathcal{B}$ will output answer 1 if and only if the answer Y of $\mathcal{A}$ satisfies $\mathsf{De}(d, Y) \neq \perp$ and $Y \neq F(X)$. Then

$$\Pr\left[\,\mathrm{PrvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 1\,\right] = \mathbf{Adv}^{\mathrm{aut}}_{\mathrm{Garble2}[\mathsf{F}]}(\mathcal{A}, k), \tag{5}$$

where b is the challenge bit of game PrvInd. We now show that

$$\Pr\left[\,\neg\mathrm{PrvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 0\,\right] \ \leq\ 2^{-k} \ . \tag{6}$$

Subtracting Eq. (5) and Eq. (6) will give the bound claimed in the theorem. Suppose that given (F, X), adversary $\mathcal{A}$ outputs $Y = (Y_1, \ldots, Y_m)$. Let $d = (Y_1^0, Y_1^1, \ldots, Y_m^0, Y_m^1)$ and let i be the smallest integer such that $Y_i \neq Y_i^{y_i}$. This integer i is well-defined if $Y \neq F(X) = (Y_1^{y_1}, \ldots, Y_m^{y_m})$. Consequently, if $\mathsf{De}(d, Y) \neq \perp$ and $Y \neq F(X)$ then Y_i must be $Y_i^{\overline{y_i}}$. This uses the fact that $Y_i^{\overline{y_i}} \neq Y_i^{y_i}$, which is true because the construction makes their type-bits unequal. Thus we have

$$\Pr\left[\,\neg\mathrm{PrvInd}^{\mathcal{A}}_{\mathcal{G},\Phi}(k) \mid b = 0\,\right] \ \leq\ \Pr\left[\,Y_i = Y_i^{\overline{y_i}} \mid Y \neq F(X)\,\right] \ . \tag{7}$$

Let $\mathtt{A}\|\mathtt{a} \leftarrow Y_i^{\overline{y_i}}$, and let $j = i + n + q - m$. Since the output bit at wire j is always y_i, during the garbling process of f', in line 205 of Fig. 10, we never encrypt token $Y_i^{\overline{y_i}}$. The string $\mathtt{A}$ is therefore independent of (F, X), and thus the right-hand side of Eq. (7) is at most 2^{-k}. $\square$

7 Applications

We believe that most applications that employ garbled circuits can be recast to use an arbitrary garbling scheme possessing one or more of the security properties we've defined. While a complete reworking of all existing garbled-circuit-using applications is beyond the scope of this paper, we sketch two examples.

First we consider the classical use of garbled circuits for two-party SFE (Secure Function Evaluation) and PFE (Private Function Evaluation). Then we consider their more recent use for securely encrypting key-dependent messages.

7.1 Two-party SFE and PFE

The classic methods for SFE and PFE combine the garbled-circuit technique with oblivious transfer (OT). The construction and proof are monolithic and complex, incorporating proofs of the garbled circuit technique. Here we aim to show how use of our formalization can simplify this process to produce proofs that are modular and thus simpler and more rigorous. We build SFE and PFE protocols by a modular combination of an arbitrary garbling scheme and an arbitrary OT protocol, reducing the security of the constructed protocol to the security of its two constituents. Besides simplicity we gain in flexibility of instantiation, for we can plug in any garbling scheme meeting our definitions and immediately get new SFE or PFE protocols that inherit the efficiency of the garbling scheme.

Classically, in SFE, a function f is public and the interaction results in party 1 learning $f(x_1 \| x_2)$ (but no more) while party 2 learns nothing, where x_i is the private input of party $i \in \{1,2\}$. In PFE, party 1 has a string x, party 2 has a function f and the outcome of the protocol is that party 1 learns $f(x)$ (but no more) while party 2 learns nothing. However, through the use of universal circuits, the two versions of the problem are equivalent. Thus, we will treat only one. We pick PFE because it is more directly obtained via garbling schemes.

It is part of our thesis that this type of program can and should be carried out rigorously and fully and that our formalization of garbling schemes enables one to do this. To this end we provide self-contained definitions of security for PFE (OT as a special case). These definitions are not the only possible ones, nor necessarily the strongest, but we need to pin something down to provide a full treatment. The setting here is that of honest but curious adversaries.

Two-party protocols. We view a two-party protocol as specified by a pair $\Pi = (\Pi_1, \Pi_2)$ of PT algorithms. Party $i \in \{1,2\}$ will run Π_1 on its current state and the incoming message from the other party to produce an outgoing message, a local output, and a decision to halt or continue. The initial state of party i consists of the unary encoding 1^k of the security parameter $k \in \mathbb{N}$ and the (private) input I_i of this party, and the interaction continues until both parties halt. We will not further formalize this process since the details are not important to what we do. What is important is that we are able to define the PT algorithm View^i_Π that on input $(1^k, I_1, I_2)$ returns the view of party i in an execution of Π with security parameter k and inputs I_1, I_2 for the two parties, respectively. Specifically, the algorithm picks at random coins ω_1, ω_2, executes the interaction between the parties as determined by Π with the initial state and coins of party $j \in \{1,2\}$ being $(1^k, I_j)$ and ω_j respectively, and returns $(conv, \omega_i)$ where the conversation $conv$ is the sequence of messages exchanged. We let $\mathsf{Out}^i_\Pi(1^k, I_1, I_2)$ return the local output of party i at the end of the protocol. This is a deterministic function of $\mathsf{View}^i_\Pi(1^k, I_1, I_2)$.

PFE. Party 1 has a string x and party 2 has a function f. The outcome of the protocol should be that party 1 learns $f(x)$. Security requires that party 2 learns nothing about x (beyond its length) and party 1 learns nothing about f (beyond side information we are willing to leak, such as the number of gates in the circuit f).

Formally a private function evaluation (PFE) protocol is a tuple $\mathcal{F} = (\Pi, \mathsf{ev})$ where Π is a 2-party protocol as above and ev is just like in a garbling scheme, meaning a PT deterministic map that associates to any string f a function $\mathsf{ev}(f, \cdot) \colon \{0,1\}^{f.n} \to \{0,1\}^{f.m}$. The correctness requirement is that for all f and all $x \in \{0,1\}^{f.n}$ we have

$$\Pr[\mathsf{Out}^1_\Pi(1^k, x, f) = \mathsf{ev}(f, x)] \;=\; 1 \;.$$

proc GetView(x, f)	Game $\mathrm{PfeSim}_{\mathcal{F},i,\Phi,\mathcal{S}}$
$b \twoheadleftarrow \{0,1\}$	
if $x \notin \{0,1\}^{f.n}$ **then return** $\perp$	
if $b = 1$ **then return** $view \leftarrow \mathsf{View}^i_{\Pi}(1^k, x, f)$	
if $i = 1$ **then return** $view \leftarrow \mathcal{S}(1^k, x, \mathsf{ev}(f, x), \Phi(f))$	
if $i = 2$ **then return** $view \leftarrow \mathcal{S}(1^k, f, \lvert x\rvert)$	

Fig. 11. Game for defining the pfe.sim security of a PFE scheme $\mathcal{F} = (\Pi, \mathsf{ev})$. Procedure Finalize$(b')$ returns $(b = b')$. The game depends on a security parameter $k \in \mathbb{N}$.

The security notion we consider is privacy in the honest-but-curious setting, meaning the parties follow the protocol and the intent is that their views do not allow the computation of any undesired information. An adversary $\mathcal{B}$ is allowed a single GetView query in game $\mathrm{PfeSim}_{\mathcal{F},i,\Phi,\mathcal{S}}$ of Fig. 11, and its advantage is

$$\mathbf{Adv}^{\mathrm{pfe.sim},\Phi,\mathcal{S}}_{\mathcal{F},i}(\mathcal{B}, k) \;=\; 2\Pr[\mathrm{PfeSim}^{\mathcal{B}}_{\mathcal{F},i,\Phi,\mathcal{S}}(k)] - 1\ .$$

We say that $\mathcal{F}$ is pfe.sim relative to Φ if for each $i \in \{0,1\}$ and each PT adversary $\mathcal{B}$ there is a PT simulator $\mathcal{S}$ such that the function $\mathbf{Adv}^{\mathrm{pfe.sim},\Phi,\mathcal{S}}_{\mathcal{F},i}(\mathcal{B}, \cdot)$ is negligible.

Oblivious transfer. The construction will utilize a protocol for 1-out-of-2 oblivious transfer where party 1 has a selection bit s, party 2 has inputs X^0, X^1, and the result is that party 1 gets X^s while party 2 gets nothing. It is convenient to assume an extension where party 1 has bits $x_1, \ldots, x_n$, party 2 has inputs $X^0_1, X^1_1, \ldots, X^0_n, X^1_n$, and the result is that party 1 gets $X^{x_1}_1, \ldots, X^{x_n}_n$ while party 2 gets nothing. Such an extended protocol may be produced by sequential repetition of the basic protocol. Formally an OT protocol is a PFE scheme $\mathcal{OT} = (\Pi^{\mathrm{ot}}, \mathsf{ev}^{\mathrm{ot}})$ where Π^{ot} is a 2-party protocol and $\mathsf{ev}^{\mathrm{ot}}((X^0_1, X^1_1, \ldots, X^0_n, X^1_n), x) = (X^{x_1}_1, \ldots, X^{x_n}_n)$. Here, a function is described by a vector $(X^0_1, X^1_1, \ldots, X^0_n, X^1_n)$, and its evaluation on an n-bit input x is $(X^{x_1}_1, \ldots, X^{x_n}_n)$. We assume a pfe.sim-secure scheme $\mathcal{OT} = (\Pi^{\mathrm{ot}}, \mathsf{ev}^{\mathrm{ot}})$ relative to the side information function $\Phi_{\mathrm{ot}}((X^0_1, X^1_1, \ldots, X^0_n, X^1_n)) = (|X^0_1|, |X^1_1|, \ldots, |X^0_n|, |X^1_n|)$.

The protocol. Let $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ be a projective garbling scheme that is prv.sim-secure over Φ. We define a PFE scheme $\mathcal{F} = (\Pi, \mathsf{ev})$ which allows the secure computation of exactly the class of functions $\{\mathsf{ev}(f, \cdot)\colon\ f \in \{0,1\}^*\}$ that $\mathcal{G}$ can garble. Party 2, on inputs $1^k, f$, begins by letting $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and parsing e as $(X^0_1, X^1_1, \ldots, X^0_n, X^1_n) \leftarrow e$. It sends F, d to party 1. Now the parties execute the OT protocol with party 1 having selection string x and party 2 having inputs $(X^0_1, X^1_1, \ldots, X^0_n, X^1_n)$. As a result, party 1 obtains $X = (X^{x_1}_1, \ldots, X^{x_n}_n)$. It now outputs $y \leftarrow \mathsf{De}(d, \mathsf{Ev}(F, X))$ and halts.

Theorem 12. Assume $\mathcal{G} = (\mathsf{Gb}, \mathsf{En}, \mathsf{De}, \mathsf{Ev}, \mathsf{ev})$ is a projective garbling scheme that is prv.sim-secure over Φ. Assume $\mathcal{OT} = (\Pi^{\mathrm{ot}}, \mathsf{ev}^{\mathrm{ot}})$ is a OT protocol that is pfe.sim-secure relative to Φ_{ot}. Let $\mathcal{F} = (\Pi, \mathsf{ev})$ be the pfe scheme constructed above. Then $\mathcal{F}$ is pfe.sim-secure relative to Φ. ☐

Proof (Theorem 12). Let $i \in \{1,2\}$ and let $\mathcal{B}$ be a PT adversary attacking $\mathcal{F}$. We build a PT adversary $\mathcal{B}_{\mathcal{G}}$ attacking $\mathcal{G}$ and a PT adversary $\mathcal{B}_{\mathcal{OT}}$ attacking $\mathcal{OT}$. By assumption, these have simulators, respectively $\mathcal{S}_{\mathcal{G}}, \mathcal{S}_{\mathcal{OT}}$. We then use these simulators to build a simulator $\mathcal{S}$ for $\mathcal{B}$ such that for every $k \in \mathbb{N}$ we have

$$\mathbf{Adv}^{\mathrm{pfe.sim},\Phi,\mathcal{S}}_{\mathcal{F},i}(\mathcal{B}, k) \leqq \mathbf{Adv}^{\mathrm{prv.sim},\,\Phi,\mathcal{S}_{\mathcal{G}}}_{\mathcal{G}}(\mathcal{B}_{\mathcal{G}}, k) + \mathbf{Adv}^{\mathrm{pfe.sim},\Phi_{\mathrm{ot}},\mathcal{S}_{\mathcal{OT}}}_{\mathcal{OT},i}(\mathcal{B}_{\mathcal{OT}}, k)\ .$$

This yields the desired conclusion. We now proceed to the constructions and analyses. We consider separately the cases $i = 1$ and $i = 2$, beginning with the former.

Adversary $\mathcal{B}_{\mathcal{G}}(1^k)$ runs $\mathcal{B}(1^k)$ to get its GetView query x, f. It will compute and return a reply *view* to this query as follows. Adversary $\mathcal{B}_{\mathcal{G}}$ queries its Garble oracle with f, x to get back $(F, X_1, \ldots, X_n), d)$. It records (F, d) as the first message in *conv*. (This message is from party 2 to party 1.) Now, for $i = 1, \ldots, n$ it lets $X^{x_i}_i \leftarrow X_i$ and $X^{1-x_i}_i \twoheadleftarrow \{0,1\}^{|X_i|}$. It then lets $view^{\mathrm{ot}} \leftarrow \mathsf{View}^1_{\Pi^{\mathrm{ot}}}(1^k, x, (X^0_1, X^1_1, \ldots, X^0_n, X^1_n))$. It obtains this by direct execution of 2-party protocol $\Pi^{\rm ot}$ on inputs $(X_1^0, X_1^1, \ldots, X_n^0, X_n^1)$ for party 2 and x for party 1. Parsing $\mathit{view}^{\rm ot}$ as $(\mathit{conv}^{\rm ot}, \omega_1^{\rm ot})$, it appends $\mathit{conv}^{\rm ot}$ to conv and then returns $\mathit{view} = (\mathit{conv}, \omega_1^{\rm ot})$ as the answer to $\mathcal{B}$'s query. Adversary $\mathcal{B}$ now outputs a bit b', and $\mathcal{B}$ adopts this as its own output as well.

Adversary $\mathcal{B}_{\mathcal{OT}}(1^k)$ runs $\mathcal{B}(1^k)$ to get its GETVIEW query x, f. It will compute and return a reply view to this query as follows. Adversary $\mathcal{B}_{\mathcal{OT}}$ lets $(F, e, d) \leftarrow \mathsf{Gb}(1^k, f)$ and parses e as $(X_1^0, X_1^1, \ldots, X_n^0, X_n^1) \leftarrow e$. It records (F, d) as the first message in conv. It makes query $\mathit{view}^{\rm ot} \leftarrow \text{GETVIEW}(x, (X_1^0, X_1^1, \ldots, X_n^0, X_n^1))$. Parsing $\mathit{view}^{\rm ot}$ as $(\mathit{conv}^{\rm ot}, \omega_1^{\rm ot})$, it appends $\mathit{conv}^{\rm ot}$ to conv and then returns $\mathit{view} = (\mathit{conv}, \omega_1^{\rm ot})$ as the answer to $\mathcal{B}$'s query. Adversary $\mathcal{B}$ now outputs a bit b', and $\mathcal{B}$ adopts this as its own output as well.

By assumption, the two adversaries we have just built have simulators, respectively $\mathcal{S}_{\mathcal{G}}, \mathcal{S}_{\mathcal{OT}}$. We define simulator $\mathcal{S}$ for $\mathcal{B}$. On input $1^k, x, y, \phi$ it lets $(F, (X_1, \ldots, X_n), d) \leftarrow \mathcal{S}_{\mathcal{G}}(1^k, y, \phi)$ and records (F, d) as the first message in conv. It lets $\mathit{view}^{\rm ot} \leftarrow \mathcal{S}_{\mathcal{OT}}(1^k, x, (X_1, \ldots, X_n), (|X_1|, |X_1|, \ldots, |X_n|, |X_n|))$. Parsing $\mathit{view}^{\rm ot}$ as $(\mathit{conv}^{\rm ot}, \omega_1^{\rm ot})$, it appends $\mathit{conv}^{\rm ot}$ to conv and then returns $\mathit{view} = (\mathit{conv}, \omega_1^{\rm ot})$.

The case $i = 2$ is much easier because party 2 obtains nothing from party 1 besides what it gets from the execution of the OT protocol and thus security follows directly from the assumption that the OT protocol is secure. $\square$

7.2 KDM-secure encryption

To be written.

Acknowledgments

Many thanks to the NSF, who sponsored this work under CNS 0904380.

References

1. M. Abadi and J. Feigenbaum. Secure circuit evaluation. *Journal of Cryptology*, 2(1):1–12, 1990.
2. B. Applebaum. Key-dependent message security: Generic amplification and completeness. In K. G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 527–546. Springer, May 2011.
3. B. Applebaum, Y. Ishai, and E. Kushilevitz. Computationally private randomizing polynomials and their applications. *Computational Complexity*, 15(2):115–162, 2006.
4. B. Applebaum, Y. Ishai, and E. Kushilevitz. Cryptography in NC^0. *SIAM J. Comput.*, 36(4):845–888, 2006.
5. B. Applebaum, Y. Ishai, and E. Kushilevitz. From secrecy to soundness: Efficient verification via secure computation. In S. Abramsky, C. Gavoille, C. Kirchner, F. Meyer auf der Heide, and P. G. Spirakis, editors, *ICALP 2010, Part I*, volume 6198 of *LNCS*, pages 152–163. Springer, July 2010.
6. B. Applebaum, Y. Ishai, and E. Kushilevitz. How to garble arithmetic circuits. In R. Ostrovsky, editor, *52nd FOCS*, pages 120–129. IEEE Computer Society Press, Oct. 2011.
7. Y. Aumann and Y. Lindell. Security against covert adversaries: Efficient protocols for realistic adversaries. In S. P. Vadhan, editor, *TCC 2007*, volume 4392 of *LNCS*, pages 137–156. Springer, Feb. 2007.
8. B. Barak, I. Haitner, D. Hofheinz, and Y. Ishai. Bounded key-dependent message security. In H. Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 423–444. Springer, May 2010.
9. M. Barni, P. Failla, V. Kolesnikov, R. Lazzeretti, A.-R. Sadeghi, and T. Schneider. Secure evaluation of private linear branching programs with medical applications. In M. Backes and P. Ning, editors, *ESORICS 2009*, volume 5789 of *LNCS*, pages 424–439. Springer, Sept. 2009.
10. D. Beaver, S. Micali, and P. Rogaway. The round complexity of secure protocols. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 503–513. ACM, 1990.
11. M. Bellare and P. Rogaway. The security of triple encryption and a framework for code-based game-playing proofs. In S. Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 409–426. Springer, May / June 2006.
12. C. Cachin, J. Camenisch, J. Kilian, and J. Müller. One-round secure computation and secure autonomous mobile agents. In *27th Intl. Colloquium on Automata, Languages, and Programming — ICALP 2000*, pages 512–523. Springer, 2000.

13. M. Chase and S. Kamara. Structured encryption and controlled disclosure. In M. Abe, editor, *ASIACRYPT 2010*, volume 6477 of *LNCS*, pages 577–594. Springer, Dec. 2010.
14. U. Feige, J. Kilian, and M. Naor. A minimal model for secure computation (extended abstract). In *26th ACM STOC*, pages 554–563. ACM Press, May 1994.
15. K. Frikken, M. Atallah, and C. Zhang. Privacy-preserving credit checking. In *Proceedings of the 6th ACM conference on Electronic commerce*, pages 147–154. ACM, 2005.
16. R. Gennaro, C. Gentry, and B. Parno. Non-interactive verifiable computing: Outsourcing computation to untrusted workers. In T. Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 465–482. Springer, Aug. 2010.
17. O. Goldreich. Cryptography and cryptographic protocols. Manuscript, June 9 2001.
18. O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game, or a completeness theorem for protocols with honest majority. In A. Aho, editor, *19th ACM STOC*, pages 218–229. ACM Press, May 1987.
19. S. Goldwasser, Y. T. Kalai, and G. N. Rothblum. One-time programs. In D. Wagner, editor, *CRYPTO 2008*, volume 5157 of *LNCS*, pages 39–56. Springer, Aug. 2008.
20. V. Goyal, P. Mohassel, and A. Smith. Efficient two party and multi party computation against covert adversaries. In N. P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 289–306. Springer, Apr. 2008.
21. A. Herzberg and H. Shulman. Secure guaranteed computation. Cryptology ePrint Archive, Report 2010/449, 2010.
22. Y. Huang, D. Evans, J. Katz, and L. Malka. Faster secure two-party computation using garbled circuits. In *USENIX Security Symposium*, 2011.
23. Y. Huang, C.-H. Shen, D. Evans, J. Katz, and A. Shelat. Efficient secure computation with garbled circuits. In S. Jajodia and C. Mazumdar, editors, *ICISS*, volume 7093 of *Lecture Notes in Computer Science*, pages 28–48. Springer, 2011.
24. Y. Ishai and E. Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *41st FOCS*, pages 294–304. IEEE Computer Society Press, Nov. 2000.
25. Y. Ishai and E. Kushilevitz. Perfect constant-round secure computation via perfect randomizing polynomials. In P. Widmayer, F. T. Ruiz, R. M. Bueno, M. Hennessy, S. Eidenbenz, and R. Conejo, editors, *ICALP*, volume 2380 of *Lecture Notes in Computer Science*, pages 244–256. Springer, 2002.
26. Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai. Cryptography with constant computational overhead. In R. E. Ladner and C. Dwork, editors, *40th ACM STOC*, pages 433–442. ACM Press, May 2008.
27. S. Kamara, P. Mohassel, and M. Raykova. Outsourcing multi-party computation. Cryptology ePrint report 2011/272, October 25 2011.
28. S. Kamara and L. Wei. Special-purpose garbled circuits. Unpublished manuscript.
29. J. Katz and R. Ostrovsky. Round-optimal secure two-party computation. In M. Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 335–354. Springer, Aug. 2004.
30. V. Kolesnikov and T. Schneider. Improved garbled circuit: Free XOR gates and applications. In L. Aceto, I. Damgård, L. A. Goldberg, M. M. Halldórsson, A. Ingólfsdóttir, and I. Walukiewicz, editors, *ICALP 2008, Part II*, volume 5126 of *LNCS*, pages 486–498. Springer, July 2008.
31. V. Kolesnikov and T. Schneider. A practical universal circuit construction and secure evaluation of private functions. In G. Tsudik, editor, *FC 2008*, volume 5143 of *LNCS*, pages 83–97. Springer, Jan. 2008.
32. B. Kreuter, A. Shelat, and C. hao Shen. Towards billion-gate secure computation with malicious adversaries. Cryptology ePrint Archive, Report 2012/179, 2012.
33. L. Kruger, S. Jha, E.-J. Goh, and D. Boneh. Secure function evaluation with ordered binary decision diagrams. In A. Juels, R. N. Wright, and S. Vimercati, editors, *ACM CCS 06*, pages 410–420. ACM Press, Oct. / Nov. 2006.
34. Y. Lindell and B. Pinkas. A proof of Yao's protocol for secure two-party computation. *Electronic Colloquium on Computational Complexity* (ECCC), TR04-063, 2004.
35. Y. Lindell and B. Pinkas. An efficient protocol for secure two-party computation in the presence of malicious adversaries. In M. Naor, editor, *EUROCRYPT 2007*, volume 4515 of *LNCS*, pages 52–78. Springer, May 2007.
36. Y. Lindell and B. Pinkas. A proof of security of Yao's protocol for two-party computation. *Journal of Cryptology*, 22(2):161–188, Apr. 2009.
37. Y. Lindell and B. Pinkas. Secure two-party computation via cut-and-choose oblivious transfer. In Y. Ishai, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 329–346. Springer, Mar. 2011.
38. D. Malkhi, N. Nisan, B. Pinkas, and Y. Sella. Fairplay — a secure two-party computation system. In *Proceedings of the 13th conference on USENIX Security Symposium-Volume 13*, pages 20–20. USENIX Association, 2004.
39. P. Mohassel and M. Franklin. Efficiency tradeoffs for malicious two-party computation. In M. Yung, Y. Dodis, A. Kiayias, and T. Malkin, editors, *PKC 2006*, volume 3958 of *LNCS*, pages 458–473. Springer, Apr. 2006.
40. M. Naor and K. Nissim. Communication preserving protocols for secure function evaluation. In *33rd ACM STOC*, pages 590–599. ACM Press, July 2001.
41. M. Naor, B. Pinkas, and R. Sumner. Privacy preserving auctions and mechanism design. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 129–139. ACM, 1999.
42. A. Paus, A.-R. Sadeghi, and T. Schneider. Practical secure evaluation of semi-private functions. In M. Abdalla, D. Pointcheval, P.-A. Fouque, and D. Vergnaud, editors, *ACNS 09*, volume 5536 of *LNCS*, pages 89–106. Springer, June 2009.

43. K. Pietrzak. A leakage-resilient mode of operation. In A. Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 462–482. Springer, Apr. 2009.
44. B. Pinkas. Cryptographic techniques for privacy-preserving data mining. *ACM SIGKDD Explorations Newsletter*, 4(2):12–19, 2002.
45. B. Pinkas, T. Schneider, N. P. Smart, and S. C. Williams. Secure two-party computation is practical. In M. Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 250–267. Springer, Dec. 2009.
46. P. Rogaway. The round complexity of secure protocols. MIT Ph.D. Thesis, 1991.
47. A. Sahai and H. Seyalioglu. Worry-free encryption: functional encryption with public keys. In E. Al-Shaer, A. D. Keromytis, and V. Shmatikov, editors, *ACM CCS 10*, pages 463–472. ACM Press, Oct. 2010.
48. T. Schneider. Practical secure function evaluation. Master's thesis, University of Erlangen-Nuremberg, 2008.
49. S. Tate and K. Xu. On garbled circuits and constant round secure function evaluation. Technical report, Computer Privacy and Security Lab, Department of Computer Science, University of North Texas, 2003.
50. J. R. Troncoso-Pastoriza, S. Katzenbeisser, and M. Celik. Privacy preserving error resilient dna searching through oblivious automata. In P. Ning, S. D. C. di Vimercati, and P. F. Syverson, editors, *ACM CCS 07*, pages 519–528. ACM Press, Oct. 2007.
51. L. Valiant. Universal circuits (preliminary report). In *Proceedings of the eighth annual ACM symposium on Theory of computing*, pages 196–203. ACM, 1976.
52. A. Yao. How to generate and exchange secrets. In *Foundations of Computer Science, 1986., 27th Annual Symposium on*, pages 162–167. IEEE, 1986.
53. A. C. Yao. Protocols for secure computations. In *23rd FOCS*, pages 160–164. IEEE Computer Society Press, Nov. 1982.

A Related Work

We do not attempt a comprehensive review of the literature (easily a monograph-length undertaking), but elaborate on some selected prior work.

Randomized encodings. Loosely related to garbling schemes, *randomized encodings* (initially *randomized polynomials*) begin with Ishai and Kushilevitz [24] and continue, with many definitional variants, in work by Applebaum, Ishai, Kushilevitz, and others [2–6, 25, 26, 47]. The authors employ language like the following [3]: function $F(\cdot,\cdot)$ is a *randomized encoding* of $f(\cdot)$ if: (correctness) there's a PT algorithm De such that $\mathsf{De}(F(x,r)) = f(x)$ for almost all r; and (privacy) there's a PT algorithm Sim such that ensembles $F(x,\cdot)$ and $\mathsf{Sim}(f(x))$ are computationally indistinguishable. To be useful, encodings must have some extra properties,[15] for example, that every bit of $F(x,r)$ depends on at most one bit of x, a property that has been called *decomposability* [26]. Proven realizations meeting these requirements [3, 4] do not closely resemble conventional realizations of garbled circuits [36, 41].

There is a large gap, even syntactically, between the notion just given and a garbling scheme. Above, no language is provided to speak of the algorithm that transforms f to F; in contrast, the thing doing this transformation is at the center of a garbling scheme. Likewise absent from the syntax of randomized encodings is anything to speak to the representation of functions; for garbling schemes, representations are explicit and central. Finally, the syntax, unlike that of a garbling scheme, does not separate the garbling of a function and the creation of a garbled input, and indeed there is nothing corresponding to the latter, the same input x being fed to f or F. The minimalist syntax of randomized encodings works well for some theory-centric applications, but does not allow one to speak of obliviousness and authenticity, to investigate the low-level efficiency of different garbling schemes, and to architect schemes to useful-in-practice abstraction boundaries.

Given the variety of related definitions, let us sketch another, the *decomposable randomized encodings* defined and used by Sahai and Seyalioglu [47]. (Despite identical names, this definition is different from that above, and different again from the decomposable randomized encodings of [26], say). The object of interest can be regarded as a pair of PT algorithms $(\mathsf{En}, \mathsf{De})$ where En maps the encoding of a boolean circuit $f\colon \{0,1\}^n \to \{0,1\}^m$ to a vector of strings $(X_1^0, X_1^1, \ldots, X_m^0, X_m^1) \leftarrow \mathsf{En}(1^k, f)$ for which decoding

[15] Otherwise, the definition is trivially met by setting $F(x,r) = f(x)$ and $\mathsf{De}(y) = \mathsf{Sim}(y) = y$.

algorithm $\mathsf{De}(X_1^{x_1}, \ldots, X_m^{x_m})$ returns $f(x_1 \cdots x_n)$. The authors demand a PPT algorithm Sim for which the ensemble of $(X_1^{x_1}, \ldots, X_m^{x_m})$ tuples induced by $\mathsf{En}(1^k, f)$ and x is computationally indistinguishable from $\mathsf{Sim}(1^k, n, |f|, f(x))$. Translating to our language, one has effectively assumed a projective scheme, a boolean circuit as input, and prv.sim security over Φ_{size}. The garbled function itself has been abstracted out of existence (in a realization, it would be dropped in the X_i^j values). Compared to a garbling scheme, one might note the lack of representation independence, granularity inadequate to speak of obliviousness, authenticity, garbled inputs, and low-level efficiency. The syntax can't handle the dynamic setting, where the adversary receives the garbled circuit before it specifies the input.

Obliviousness and authenticity. Some prior papers exploit obliviousness and authenticity of garbled circuits to achieve desired applications: private medical diagnostics [9], verifiable computation and private verifiable computation [16], and correctable verifiable computation [5]. The notions are not seen as properties of a stand-alone primitive corresponding to a garbling scheme.

In the last of the works mentioned, Applebaum, Ishai, Kushilevitz [5] describe the following generic transformations from privacy to obliviousness and to authenticity. (1) Obliviousness: instead of garbling a circuit f, choose $r \twoheadleftarrow \{0,1\}^m$ and garble a circuit g such that $g(x) = f(x) \oplus r$ for every $x \in \{0,1\}^n$, where $n = f.n$ and $m = f.m$. (2) Authenticity: instead of garbling a circuit f, choose $K \twoheadleftarrow \{0,1\}^m$ and garble a circuit g such that $g(x) = f(x) \parallel \mathrm{MAC}_K(f(x))$ for any $x \in \{0,1\}^n$. Applied to Garble1, the transformations lead to schemes slightly (for (1)) or substantially (for (2)) less efficient that Garble2; and (2) requires a cryptographic assumption. More fundamentally, Applebaum *et al.* do not formalize any definition for the obliviousness or authenticity of a garbling scheme.

The only work that explicitly defines obliviousness and authenticity in this domain is a recent paper of Kamara, Mohassel, and Rakova [27]. Still, their syntax is designed specifically for their application; for example, a circuit's input is a pair (x_1, x_2), a garbled circuit's input is (X_1, X_2), and the encoding function takes an input x and an index $i \in \{1, 2\}$ and outputs the corresponding X_i. Their notion of obliviousness requires hiding only the input, while obv.ind and obv.sim require one to hide both the input and the function.

Obscuring topology. We are not the first to observe that conventional means to garble a circuit obscure each gate's function but not its topology. A 2002 paper of Pinkas [44, Section 2.3] already remarks that "In this form the representation reveals nothing but the wiring of the circuit". Later, Paus, Sadeghi, and Schneider [42] use the phrase "circuit topology" to name that which is revealed by conventional garbled circuits. Nevertheless, the topology of a circuit is never formalized, and nobody ever proves that that some particular scheme reveals only the topology. We are also the first to explain the equivalence between the prv.sim and prv.ind notions relative to Φ_{topo}.

Eclectic representations. Scattered through the literature one finds computational objects other than boolean circuits that are being garbled; examples include arithmetic circuits [6], branching programs [9], circuits with lookup tables [40], DFAs [50], and ordered binary decision diagrams [33]. The range suggests, to us, that general-purpose definitions for garbling schemes ought not be tied to circuits.

Concurrent work. Concurrent work by Kamara and Wei (henceforth KW) investigates the garbling of *structured circuits* [28], a computational model they put forward resembling ordinary circuits except that gates perform operations on an arbitrary data structure. As part of this work, KW define what they too call a garbling scheme. Their syntax is similar to ours, but without the function ev. Over this syntax KW define Ind1 and Sim1 security. These notions, unlike ours, ask only for input-hiding, not function hiding. They show these definitions are equivalent for *sampleable* circuits. KW go on to give dynamic versions of their definitions, Ind2 and Sim2, and an unforgeability notion, Unf2. These definitions resemble the weaker form of the dynamic-security definitions (prv.ind!, prv.sim!, and aut!) mentioned in our Introduction and the subject of separate work.

Although KW speak of circuits as finitary objects described by DAGs, they appear to have in mind families of circuits, indexed by a security parameter (otherwise, we do not know how to make sense of samplability, or phrases like *polynomial size circuits*). Unlike our treatment, circuits are not provided by the adversary; security notions are with respect to a given circuit. A garbling scheme is provided in KW, but not a "conventional" one: it garbles a structured circuit and is based on a collection of *structured encryption schemes*, a notion from Chase and Kamara [13]. For the protocol to make sense with respect to the definitions given, the latter should be reinterpreted as applying to structured circuits.

B Universal Circuits

An (n, q)-*universal circuit* is a circuit $\mathscr{U}$ having q distinguished gates $g_1, \ldots g_q$ such that:

- It takes two inputs f and x where $|x| = n$ and f is the encoding of a circuit of input length n and at most q gates.
- For any input (f, x), when we evaluate $\mathscr{U}$ on (f, x), the bit obtained at the outgoing wire of g_i is exactly the bit obtained at the outgoing wire of gate i of f when we evaluate f on x.

A universal circuit must have size $\Omega(q \log q)$ because, by a counting argument, there are $\Omega(q2^q)$ circuits of q gates. Valiant [51] designs an (n, q)-universal circuit of fanout 4 and size $19(2q+m)\lg(2q+m)+9q$, which is asymptotically optimal, where m is the number of outputs of the original circuit. There are other constructions [31, 48], but their asymptotic size is bigger.