Collusion-Preserving Computation without a Mediator

Michele Ciampi
University of Edinburgh
michele.ciampi@ed.ac.uk

Yun Lu
University of Edinburgh
Y.Lu-59@sms.ed.ac.uk

Vassilis Zikas
Purdue University
vzikas@cs.purdue.edu

***Abstract*—Collusion-free (CF) and collusion-preserving (CP) protocols enrich the standard security offered by multi-party computation (MPC), to tackle settings where subliminal communication is undesirable. However, all existing solutions make arguably unrealistic assumptions on setups, such as physical presence of the parties, access to physical envelopes, or extreme isolation, where the only means of communication is a star-topology network. The above state of affairs remained a limitation of such protocols, which was even reinforced by impossibility results. Thus, for years, it has been unclear if and how the above setup assumptions could be relaxed towards more realistic scenarios. Motivated also by the increasing interest in using hardware tokens for cryptographic applications, in this work we provide the first solution to collusion preserving computation which uses weaker and more common assumptions than the state of the art, i.e., an authenticated broadcast functionality and access to honestly generated trusted hardware tokens. We prove that our protocol is collusion-preserving (in short, *CP*) secure as long as no parties abort. In the case of an aborting adversary, our protocol still achieves standard (G)UC security with identifiable (and unanimous) abort.**

Leveraging the above identifiability property, we augment our protocol with a penalization scheme which ensures that it is not profitable to abort, thereby obtaining CP security against incentive-driven attackers. To define (and prove) this latter result, we combine the Rational Protocol Design (RPD) methodology by Garay *et al.* [FOCS 2013] with the CP framework of Alwen *et al.* [CRYPTO 2012] to derive a definition of security in the presence of incentive-driven local adversaries which can be of independent interest. Similar to existing CP/CF solutions, our protocol preserves, as a fallback, security against monolithic adversaries, even when the setup (i.e., the hardware tokens) is compromised or corrupted. In addition, our fallback solution achieves identifiable and unanimous abort, which we prove are impossible in previous CP solutions.

I. Introduction

Subliminal communication channels in protocols allow parties to embed extra information into protocol messages, often without being detected. The existence of subliminal channels is problematic in several applications of secure computation. In large-scale distributed systems, for instance, subliminal channels could allow two parties to coordinate their actions (i.e., collude) even if they may not have been aware of each other in advance. Such collusions have severe consequences in game theoretic applications, where stability, e.g., Nash equilibrium, is defined in terms of isolated strategies. An example is the prototypical application of distributed cryptography, namely, playing poker in a distributed manner [32]. An MPC protocol which allows collusions changes the rule of the game — think of playing poker against colluding opponents. In the quest to combine game theory and cryptography, a number of works [1], [4], [20], [44], [45] put forth new security notions, and in particular the notion of *collusion-freeness (CF)*. Analogous to simulation-based security, where a protocol is "secure" if the view of the adversary (controlling corrupt parties) can be emulated by a simulator, a protocol is *collusion-free*, if the view of individual corrupt parties can be emulated by individual non-colluding simulators. However, collusion freeness is impossible when parties are connected by pairwise communication channels—this is straightforward to see since such channels can directly be used for coordination.

The above led to the proposal of alternative models that enable collusion freeness. The common feature of these models is that, unlike traditional cryptographic definition that assume a worst-case monolithic adversary, these alternative models consider isolated/local adversaries.[1] The two most typical models are assuming players are physically collocated, have access to a semi-trusted "ballot box," and can communicate publicly via (physical) envelopes [44], [45], or assuming that parties are connected to a semi-trusted party (via standard communication channels) called the *mediator* [1], [4] in a star network topology. Roughly, in each round, the mediator performs a two-party computation with each party individually, in order to prevent subliminal communication between parties.

Unfortunately, collusion-freeness, as a standalone definition, is not enough to limit collusion when parties can be engaged in multiple protocols. As argued by Alwen et al. [2], unlike what one might expect, a CF protocol (secure according to the definition of [1], [4]) does not necessarily preserve its "collusion-freeness" when composed with other protocols. Alwen et al. gives the following simple counter-example. Suppose a CF protocol is augmented with the following: it allows two parties A and B to collude only if party B can provide the correct (randomly-generated at run-time) λ-bit password, but the password is given only to party A. The protocol remains CF, since B can only guess the password with negligible probability. However, if through executing another protocol party A can communicate λ-bits to B, then A can send the password and collusion-freeness is lost. This limitation motivated [2] to introduce the notion of collusion-preserving computation (CP). Intuitively, this notion can be seen as a universally composable (UC) [15] extension of CF, designed to

[1] Notably local adversaries are a generalization of the monolithic-adversary model; indeed, the latter can be captured by giving local adversaries the ability to communicate through the (assumed) network (cf. [4], [20]).

explicitly address the above issue. (An alternative model capturing collusion-preserving universally composable computation via local adversaries was concurrently and independently proposed by Canetti and Vald [20].) As a concrete example, suppose a protocol Π is CP, and suppose while executing Π, a corrupt party p_i uses an external communication channel to send another party p_j a message m (e.g. a secret value only p_i knows). CP ensures that knowing m does not allow p_j to gain *even more* information in Π, such as p_i's input or output, other than information already implied by m. Importantly, CP supports composition and it is modeled with respect to arbitrary communication resources.

Continuing the line of works in the mediated model, [2] constructed a collusion-preserving (CP) protocol, which achieves the following fallback guarantee: Assuming a global augmented common reference string (ACRS) functionality [17], when parties are arranged in a star topology with the mediator in the center, there exists a protocol which (1) CP emulates any given functionality if the mediator is honest, and (2) remains GUC secure [17]—i.e., secure with abort according to the monolithic-adversary definition—even if the mediator gets corrupted.

On the downside, the CP protocol of [2] in the mediated model presents several limitations. First, it is straightforward to prove that desirable properties such as fairness and identifiable abort are impossible in the mediated model when the mediator might get corrupted (App. C). This can lead to undesirable situations—such as a poker tournament where a player can always abort the game when he realizes he is losing, without being identified as a cheater. Second, the protocol (compiler) from [2] takes explicit steps to ensure that an abort does not allow parties to correlate their strategies, by making sure that an abort is observable only in a final round that is deterministically fixed at the beginning of the protocol. However, this means that their solution cannot be used to compute reactive functionalities, since a party can signal by means of aborting in an intermediate output (in the appendix we provide a description of a concrete collusion strategy which this allows). Lastly, the protocol is not round-efficient. Having a deterministic upper bound on the number of rounds means that, if the goal is to unanimously decide on whether or not an abort occurred even in the fallback setting where the mediator is compromised, the protocol always needs a linear number of rounds. This is true because a generic compiler for such a functionality would imply deterministic broadcast which needs linearly many rounds [24]. In fact, in the [2] compiler, each round is emulated by a round-robin sequential interaction of each party with the mediator, yielding an additional linear multiplicative blowup in the round complexity.

In this work we circumvent several of the limitations of [2] and extend its applicability towards a framework for collusion-preserving protocols over public networks. To our knowledge, this is the first work proposing a solution that breaks the deadlock of collusion-free/preserving computation, which was believed to only apply to the mediated model or require physical presence of parties in the same room. Concretely, our solution replaces the mediator by the strictly weaker (as we argue later) assumption of an authenticated broadcast channel and honestly generated hardware tokens. As we will show later (Sec. III): (1) Capturing such hardware tokens in a collusion preserving framework is non-trivial (2) Our protocols achieve CP when the adversary does not abort (and in fact CP is impossible with the broadcast channel when the adversary can abort). To disincentivize aborts, we show how to leverage the publicly identifiable abort (Section V-1) property of our protocol to concretely penalize (e.g., via the blockchain) aborting parties. The fact that our protocol is CP means that any external communication introduced by the blockchain will not lead to even more correlation in our protocol. To capture incentive-driven attackers in a composable manner, we combine and extend the CP framework with the Rational Protocol Design (RPD) methodology [28]. We believe that both our treatment of hardware tokens in CP, and our incentives model which we term RPD-CP, can be of independent interest.

As fallback in case the hardware tokens may be compromised, our proposed protocol protects the inputs of honest parties and ensures identifiable (unanimous) [2] abort while guaranteeing termination (cf. [38]). This is the analogue—in the token-hybrid setting with a broadcast channel—of the fallback property of [1], [2], [4] which preserves privacy against a corrupted mediator at the center of a a star-network. In fact, our fallback is stronger than what the mediated-model permits [2]; indeed, the star-network topology makes it impossible to obtain identifiable (unanimous) abort against a corrupted mediator. The reason is that since the mediator controls the communication, it is impossible to correctly detect if a malicious party did not sent a message, or if it is the malicious mediator is pretending that an honest party has stopped replying (App. C).

1) Overview of our contributions: We now present our contributions in more detail. The goals of our work are to construct CP protocols that (1) replace the mediator-centered star-topology network by weaker resources which are closer to modern communication networks, (2) achieve stronger fallback security properties and (3) ensure CP computation of even reactive functionalities by proving that incentive-driven attackers will not abort, in a security model that incurs the negative cost of abort to the adversary. These goals bring the theory of CP closer to capturing real world applications such as a decentralized-dealer poker game. For reference, a comparison of our results to [45] and [1], [2] can be found in Table I. In more detail, our protocol emulates CP functionalities, including those with private actions, when no abort occurs. We disincentivize aborts via a concrete penalization scheme and in addition, we achieve fallback security maintaining identifiable abort (and unanimous abort). That is, even when hardware tokens are compromised and the parties are allowed to abort we still retain standard GUC security.

[2]Recall that security with *unanimous abort* guarantees that either all or none of the honest parties receive the output while *identifiable abort* ensures any party causing an abort can be identified (and excluded from future executions).

The abort column in the table specifies the number of bits of subliminal communication possible when an abort happens.

As a first step towards our goals, we show how to construct a collusion-preserving protocol $\Pi^{\mathtt{HT}}$ allowing n parties to CP emulate any given CP-well-formed functionality—intuitively these are CP versions of well-formed functionalities [19] which, when everyone gets corrupted, give up on collusion preservation (Def. 9). Our protocol uses as resources (1) honestly generated stateful trusted hardware tokens (HTs) and (2) an authenticated broadcast channel available during protocol execution[3]. Our protocol, which can CP-emulate any functionality as long as no abort occurs, improves upon the CF-protocol of [45], which is also based on broadcast but is not CP and does not emulate games with private actions (i.e. actions that are not publicly observable). We note that our protocol, analogously to that of [45], remains (G)UC secure with identifiable abort (though not collusion free) in case of abort. We remark that $\Pi^{\mathtt{HT}}$ only requires two (broadcast) communication rounds unlike that of [45] or [2]. Furthermore, unlike the mediator from [2], tokens do not need to know in advance what computation they will be used for, and only need to be initialized with correlated randomness independent of the protocol, discussed in the overview of our techniques.

The protocol $\Pi^{\mathtt{HT}}$ only offers security guarantees when the tokens are uncompromised. This is arguably a strong assumption since in reality the adversary may attempt to break the security of the tokens. For this reason we present a protocol compiler $\Pi^{\mathtt{HT\text{-}FBS}}$ which on input a (standard (G)UC secure) protocol with *unanimous* or *identifiable abort*, outputs a protocol with the same CP guarantees as $\Pi^{\mathtt{HT}}$ and, additionally, preserves (G)UC security properties (i.e., security with *unanimous* or *identifiable abort*, respectively) of the compiled protocol, even when hardware tokens are compromised. Specifically, even if the memory/code of corrupt parties' tokens can be read/reprogrammed and the secret keys of honest parties' tokens are leaked, the protocol is still (G)UC secure. This improves upon the fallback security of [2] where, due to model idiosyncrasies, these properties are impossible to achieve when the mediator is corrupted, even with honest majority (also in this case the reason is that the mediator has full control of the network, see App. C). We note that $\Pi^{\mathtt{HT\text{-}FBS}}$, even with the extra fallback guarantee, has round complexity that is still lower than the mediated-model solution of [2].

We believe that the above corruption model for the fallback solution is quite realistic, and captures most of the attacks that have been performed on hardware tokens. For this reason, our main theorems are proven in this corruption model. However, for sake of completeness, we will also briefly sketch how to construct a protocol that preserves (G)UC security even in the case where also the tokens of the honest parties are fully compromised (i.e., the memory/code of any token can be read/reprogrammed). However, this protocol will require a higher communication complexity, and we lose the property of identifiable abort.

[3]The broadcast channel used in this work guarantees that messages are always delivered to the parties [27], [35].

	Channel, assumption	Id-abort (fallback)	U-abort (fallback)	Private actions	Abort
Lepinski et al., STOC 2005	Broadcast+physical presence of parties, physical envelopes	✓	✓	✗	$\mathsf{poly}(\lambda)$ bits
Alwen et al., Crypto 2009 **Alwen et al., Crypto 2012**	Star topology, honest mediator	✗	✗	✓	$\Omega(\log \lambda)$ bits for reactive functionalities
This work	Authenticated Broadcast, hardware tokens	✓	✓	✓	Disincentivized

TABLE I
COMPARISON WITH EXISTING APPROACHES. ID-ABORT: IDENTIFIABLE ABORT, U-ABORT: UNANIMOUS ABORT.

As an additional contribution, we combine the Rational Protocol Design (RPD) [6], [28] framework with CP and define a new model termed *RPD-CP*. Within this model we define a CP-security notion *collusion preserving attack payoff* (in short, *CPAP*), which intuitively corresponds to security against any combination of incentive-driven local adversaries. We identify a natural class of utilities under which non-aborting strategies are strictly dominant in $\Pi^{\mathtt{HT}}$ and $\Pi^{\mathtt{HT\text{-}FBS}}$. That is, these protocols are collusion-preserving according to CPAP against adversaries bounded by this utility. We believe that RPD-CP can be used to derive formal composable versions of security statements with fair-compensation [5], [11], [39]–[41] which we think is an interesting future direction. Finally we propose a concrete penalization mechanism, which may be implemented via e.g., the blockchain, that induces utilities in the above class. Combining this with the notion of CPAP, we can prove that any adversary who maximizes its revenue (or at least minimizes loss) will not abort, making our protocol CP secure against such adversaries. We note in passing that the ability of the CP security definition to make formal statements in the presence of a global blockchain demonstrates the power of the CP definition. Indeed, such a CP statement concretely treats the protocol-independent communication as an external channel; what CP ensures then is that the protocol does not increase the information communicated through this external medium. Thus, informally, if one can devise an equilibrium in an ideal poker game (with a trusted dealer), where the players can also access the blockchain, then by adapting the results from [2] we can prove that this equilibrium would be preserved when the poker dealer is replaced by our CP protocol.

2) Organization of the paper.: Sec. III: Overview of our constructions and main techniques. Sec. IV: Preliminaries. Sec. V: Protocol $\Pi^{\mathtt{HT}}$ for CP functionalities assuming tokens, broadcast and no abort. Sec. V-A: Protocol $\Pi^{\mathtt{HT\text{-}FBS}}$ with fallback GUC security against compromised tokens. Sec. VI: New framework RPD-CP for defining CPAP—security of CP protocols against rational attackers. Sec. VI-B: Proof our protocols $\Pi^{\mathtt{HT}}$ and $\Pi^{\mathtt{HT\text{-}FBS}}$ are CPAP secure. Sec. VI-C: Penalization scheme to make the utilities defined in Sec. VI concrete.

II. RELATED WORKS AND DISCUSSIONS

A. Collusion Freeness and Preservation

We extend the comparison of our work with existing results on collusion freeness (CF) and preservation (CP). The work

of Lepinski et al. [45] achieves collusion-freeness for parties that communicate with a broadcast channel, assuming access to a physical primitive called "envelopes". They motivate the use of envelopes by proving that with only (authenticated) broadcast channels, CF is impossible, even when the adversary does not cause the computation to abort. The idea is that corrupted parties can share the same random tape before the start of protocol execution. Since all messages are sent through broadcast, all corrupted parties will have the same view and thus emulate a monolithic adversary. In our work, parties also communicate via a broadcast channel, but we circumvent this impossibility since the random tape of a corrupted party is not decided by the party himself, but by a hardware token (whereas Lepinski et al. generate randomness through coin-tossing and hide the result in envelopes).

Using stateful hardware tokens, we can also circumvent another impossibility result of Lepinski et al., that CF protocols with private actions/inputs are not possible even with envelopes. The impossibility comes from the corrupted parties being able to see (different) protocol messages that different private actions generate, and choosing his private action such that the resulting messages convey subliminal information about the private action itself. However, stateful hardware tokens can be programmed to prevent users from changing his input. That is, ensure parties cannot see messages generated by different inputs. Thus, our protocol remains CP even with private actions. As described in the introduction, in the mediated model, [1], [4] achieve CF and [2] achieves CP. In [20] the authors consider the notion of UC security with local adversaries. This notion is more general than the notion of CP as it captures more security flavors and more sophisticated corruption models. In particular, the notion introduced in [20] capture the scenario where there are independent of clusters of corrupted party, whereas [2] assume that each adversary works in isolation.

1) Authenticated Broadcast Assumption: We compare authenticated broadcast with assumptions from previous works. Existing works on CF and CP either require parties to be present in person [45]—to pass around physical envelopes, or rely on restricted star-topology communication networks [1], [2]. These assumptions are used far less than the (authenticated) broadcast channel common in MPC literature, and have been criticized as overly restrictive and/or unrealistic.

In particular, the physical presence assumption of [45] makes the resulting protocols inapplicable to the standard cryptographic setting, where the protocol is played by interactive Turing machines (ITMs). Similarly, the mediated model [1], [2] requires both complete isolation of the parties and a special star-network topology where the mediator (at the center) is required to both preserve the privacy of the communication and generate (pseudo)randomness—these requirement are proven to be necessary for CP in the mediated model [2]. This is in contrast to the less demanding assumption of authenticated broadcast combined with honestly generated hardware tokens, which not only we believe is more natural—authenticated broadcast is a standard assumption in the cryptographic protocols literature and trusted hardware is a far less exotic assumption than it used to be, but is also formally weaker as discussed in Section II-C.

As in previous works, if our communication resource (i.e. broadcast) is malicious (e.g., allows undetectable communication between corrupt parties) then our protocols are no longer CP. (In fact it is easy to verify that such undetectable communication would make CP infeasible.) Nonetheless, analogous to "fallback" security in the CP setting with a corrupted mediator [2], our protocol still preserves its standard (G)UC security guarantees under a malicious broadcast resource (see Section V-A2).

B. Playing Games over the Blockchain

The question of playing games such as poker over the Internet has recently attracted considerable attention, fueled by the new capabilities introduced by smart-contract-enabled (cryptocurrency) blockchains, such as Ethereum. In a nutshell, this technology makes it possible to ensure that parties cannot avoid paying their bid amount when they lose without the need of a trusted escrow—by having them commit their bids on the blockchain, in a smart contract that releases them to anyone that presents evidence of winning. Furthermore, the same technology enables a mechanism that punishes cheating—or early aborting—by making parties commit collateral that they can only claim if evidence is presented that they completed their protocol. This method of penalization has been used by, e.g., [5], [11], [39]–[41], which gave rise to a number of proposals for decentralized poker protocols [12], [42]. However, all these works use standard multi-party computation, thus even players who do not know each other can collude via protocol messages.

C. Stateful Tamper-Resilient Hardware Tokens

Previous works (e.g. [3], [23], [25], [31], [34], [37], [47], [49]) have based (UC-)secure protocols on stateful, tamper-resilient hardware. In addition to theory, trusted hardware such as Intel's SGX ([8], [9], [26]) and Bitcoin Hardware Wallets, have also been deployed in real life. The protocol of Lepinski et al. [45] is based on physical envelopes, a kind of trusted hardware as well. They constructed CF (though not CP) protocols for games with public actions, where parties pass around such envelopes.

We extend the application of hardware tokens by presenting the *first* (in our knowledge) collusion-preserving protocol based on stateful trusted hardware tokens. We emphasize that while improving upon the limitations and impossibilities of previous works, we do not remove the need for trust completely. However, instead of relying on a trusted mediator to achieve CP, we replace it with authenticated broadcast (discussed in Section II-A1) and trusted hardware tokens. Intuitively, these assumptions are weaker than the mediator, since an honest mediator in a star topology can act as a broadcast channel and as a set of hardware tokens. In fact in App. A we prove a formal separation of the assumption of

a trusted mediator from the one used here (i.e., the combination of broadcast and honestly generated hardware tokens) demonstrating that our assumption is indeed strictly weaker than that of a trusted mediator ; in particular we show that a trusted mediator allows for CP even in the presence of an aborting adversary, which is impossible in our setting.

With trusted tokens, we improve upon the solution of [45] by achieving composition, games with private actions, and only requiring parties to broadcast messages instead of physically exchanging tokens. We also circumvent the impossibilities of previous works in collusion-preservation, as discussed in the introduction. To improve the practicality of our solutions, in Section V-A we propose protocols that provide fallback security in case of compromised tokens (we discuss two levels of severity of compromise), as well as address practical issues in implementation (Section V-2).

Below, we detail the properties of our token assumption. We follow the approach of [3] to model hardware tokens as ideal functionalities. Our tokens require the following properties:

- *Stateful*: The token has internal memory which may be read/updated.
- *Trusted and Tamper-resilient*: The token manufacturer is trusted, and no one except the token itself can read or write its contents. In Section V-A we also sketch solutions (with some trade-offs) which preserve GUC security given untrusted or non-tamper-resilient tokens.
- *Isolated* from its creator: Only the token owner can query and get the outputs generated by the token. In particular, we require that no other parties may communicate with the token.

The formalism of isolated, tamper-resilient and stateful hardware tokens was introduced by [37]. There, the creation process of a token is described by a "wrapper" functionality which allows parties to store and run (possibly several) ITMs representing the code and memory of tokens. The wrapper functionality models malicious token creators in the protocol. However, to achieve CP, we assume that all the parties hold honestly-generated tokens that share some common private information. Thus, following [3], we model each token as a ideal functionality without using the wrapper. Proving collusion-preservation based on hardware tokens presents several unique technical challenges. We detail these issues and their solutions in Section III.

III. Overview of our techniques

Let $\mathcal{P}$ be a set of parties who wish to compute a function f in a collusion-preserving way. We assume that each party $p_i \in \mathcal{P}$ has access to a hardware token (HT) $\mathtt{HT}_i$. All HT's contain as secret information pseudo-random function's (PRF) keys k_0, k_1 and a master secret key msk (a secret key for a strong signature scheme). The public interface of the hardware tokens is represented by the master public key mpk for msk. We refer to the party $p_n \in \mathcal{P}$ as the *leader*. Moreover, each execution of the protocol is uniquely identified by a session id $\mathsf{sid} \in \mathbb{N}$.

a) Collusion preserving protocol via HT and non-aborting adversary with broadcast: Roughly, our first protocol $\Pi^{\mathtt{HT}}$ works as follows: Each party $p_i \in \mathcal{P} - \{p_n\}$ sends his input, encrypted by his token $\mathtt{HT}_i$, to a designated leader party. Upon receipt, the leader gives these messages, along with his own input, to his own token. The token then computes the output, which the leader forwards to the other parties.

In more detail, each token $\mathtt{HT}_i$ uses $R_0 \leftarrow \mathsf{PRF}(k_0, \mathsf{sid})$ as randomness to generate an encryption key sk. In addition, it uses $R_1 \leftarrow \mathsf{PRF}(k_1, \mathsf{sid}||i)$[4] to generate a pair of session signing-verification $(\mathsf{sigk}_i, \mathsf{vk}_i)$ keys for a strong signature scheme and certifies them by signing $\mathsf{vk}_i||\mathsf{sid}||i$ with the master secret key msk thus obtaining cert_i.[5] We refer to the encryption key sk as the *session encryption key* and to $(\mathsf{sigk}_i, \mathsf{vk}_i)$ as the *session signing-verification keys*[6]. Note that the session encryption key sk is common to all the hardware tokens (since all the tokens share the same PRF keys).[7]

After the session keys have been generated, the hardware token $\mathtt{HT}_i$ encrypts his input x_i, signs this encrypted value together with f using sigk_i, and sends the encryption, the signature, and the certificate cert_i over the broadcast channel. The leader p_n collects all the encrypted values and signatures, and gives them to his hardware token $\mathtt{HT}_n$ along with his input x_n, the function f, and sid. His hardware token $\mathtt{HT}_n$ first checks that all certificates and signatures are valid and the inputs are consistent with the function f. Then, if all the checks are successful, $\mathtt{HT}_n$ generates, as described earlier, the session encryption key sk and decrypts all the encrypted inputs of the other parties using sk (we recall that the PRF key k_0 is shared among all the hardware tokens). Using everyone's (decrypted) inputs $x_1, \ldots, x_n$, , $\mathtt{HT}_n$ evaluates $y_1, \ldots, y_n = f(x_1, \ldots, x_n)$ and for $i = 1, \ldots, n-1$ encrypts y_i using sk and signs the (concatenation of the) encrypted values together with f using sigk_n. The encryptions and the signature uses randomness generated from evaluating $\mathsf{PRF}(k_1, \mathsf{sid}||n)$. Finally, the leader p_n propagates the output of $\mathtt{HT}_n$ on the broadcast channel. Each party p_i, upon receiving a message, forwards it to $\mathtt{HT}_i$, which verifies the certificate, the signature, and the consistency between f and sid. If the checks are successful then $\mathtt{HT}_i$ uses sk to decrypt and output y_i.

Using hardware tokens allows us to achieve the following: 1) generate fresh randomness and session keys for each new session id sid, that are hidden from the parties themselves, and 2) certify that each sid is used only once. Intuitively, for 1), the randomness used in the computation must be hidden from the parties themselves to achieve CP over broadcast. A concrete example to demonstrate why hiding the randomness is important: Suppose a party Alice has access to a (limited) external channel and uses it to send this randomness to

[4] As an abuse of notation we refer to the identity of a party p_i with i.

[5] We use $||$ as the concatenation operator.

[6] Unless otherwise specified, a signing-verification key always refers to a *session* signing-verification key.

[7] Intuitively, we use session-keys instead of the master secret key because these keys will be leaked to the simulator. Leaking the master secret key would completely compromise the token. A more detailed discussion follows later this section.

another party Bob. If, for example, the randomness is used as Alice's decryption key in the protocol, Bob now can also decrypt messages directed towards Alice, since Bob sees these encrypted messages over broadcast. This breaks CP as the limited external communication led to additional information, e.g., Alice's output, to be leaked to Bob. For 2), to restrict any sid to one-time use, our stateful hardware token stops replying when a sid is used more than once. This is necessary to prevent an adversary from using the same sid (thus the same randomness) to evaluate different inputs. Otherwise, he can send a subliminal message by picking an input where, for example, the resulting encrypted message has its first two bits equal to the first two bits of his input—breaking collusion-preservation. This adversarial strategy was indeed observed in [45], limiting the games they consider to those with publicly observable actions. The proof that that $\Pi^{\mathtt{HT}}$ is CP for non-aborting adversaries intuitively comes from the fact that $\Pi^{\mathtt{HT}}$ is deterministic given the tokens (which fixes the PRF and msk keys) and the sid we use. We note that while this may appear contradictory (i.e., to obtain a secure protocol you need entropy [33]), our protocol achieves security and collusion-preservation as the tokens generate fresh (pseudo)randomness for each sid (which is used only once). $\Pi^{\mathtt{HT}}$ also enjoys identifiable abort, and more interestingly, any external party observing the execution of the protocol without participating it can identify malicious behavior, by verifying signatures of messages on the channel. Following [39] we refer to this property as *publicly identifiable abort*. To reduce the amount of token memory required for checking that each sid is used only once, we propose two alternate solutions in Section V-2.

b) Tokens in collusion-preserving computation: One may wonder why hardware tokens cannot directly use their master secret key to authenticate protocol messages. The reason lies in the locality restrictions that the CP model places on the ideal world adversary (the simulator). Specifically, CP requires the existence of a local simulator $\mathcal{S}_i$ for each adversarial party p_i, and mandates that simulators cannot communicate with each other except if the environment explicitly allows them to. Thus, setups (in our case, tokens) which naturally introduce correlations between parties are tricky to define and use, especially when the protocol is executed over a broadcast channel. To understand the issue, one needs to observe that in a protocol over a broadcast channel, all parties expect to see exactly the same messages from this channel. Indeed, one of the novelties of our work is to show how in the real-world, i.e., in the protocol execution, the correlations embedded in the tokens can be leveraged to ensure that every (honest-protocol) broadcast message is predictable by any token. However, this correlation in the views inherently requires the simulators' views to be correlated in some way, in order to generate the same exact protocol messages. For instance, if $\mathcal{S}_1$ (the ideal world adversary with 1 as their ID) reports to the environment that he broadcasted message m in round ρ, then each $\mathcal{S}_j$ should also report to the environment that they heard this message. However, for this we need to allow the simulators to correlate their response. A naive first approach would be to allow them to interact over some underlying communication network. However, this defeats the purpose of collusion preservation, as it explicitly introduces a venue of arbitrary correlations/collusion. A second approach would be to also offer the simulators access to correlated tokens. But this leads to a new technical issue: In order to simulate the token-hybrid protocol, our simulator needs to have extra control of the hardware tokens (e.g., be able to program it). In fact, such asymmetry between the capabilities of the adversary and the simulator is proven necessary in various related settings (e.g., programmable random oracle).

We tackle the above issue by considering the hardware token as a (global) setup functionality and embedding a trapdoor inside. To ensure that only the simulators in the ideal world can use the trapdoor, we employ a technical trick inspired by [18] for the global random oracle. At a very high level, the trapdoor allows the simulators to produce the same signed messages, making their views on protocol messages consistent. Specifically, it gives access to a set of pre-computed messages which contain no information about the input of the parties. These messages are indistinguishable from the messages that would be generated in the real world and are properly signed with respect to the n signing-verification session keys as they would be in a real world execution. For example, for the protocol we have just described, the trapdoor would allow the simulator to get a set of encryptions of 0, all authenticated with respect to the corresponding signing-verification session keys. To unable such a mechanism, we introduce a token global-functionality that allows functionalities registered to it to send a special command $(\mathsf{Trapdoor}, \mathsf{sid})$. The registered functionalities can then relay the trapdoor information it receives to its simulators, allowing them to complete their simulations.

Most importantly, this trapdoor information remains useless for other protocols, preserving composability with other CP protocols. More concretely, consider an extension $\mathcal{F}^\star$ of $\mathcal{F}$, which behaves exactly as $\mathcal{F}$ but accepts an additional command GetTrapdoor from the ideal world adversary. In the simulation each simulator can send to $\mathcal{F}^\star$ the command $(\mathsf{GetTrapdoor}, \mathsf{sid})$. Upon receiving this command, $\mathcal{F}^\star$ sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to the token functionality if and only if sid is equal its session id. The token functionality, upon receiving the command $(\mathsf{Trapdoor}, \mathsf{sid})$ from $\mathcal{F}^\star$, sends to $\mathcal{F}^\star$ the trapdoor information (e.g., the authenticated encryptions of 0). When $\mathcal{F}^\star$ receives a reply from the token functionality, it is forwarded to the simulator. Note that we leak only messages that can be used within a specific session id, since they are signed using signing key that are bonded to one session id. Indeed, as discussed previously, for each session we create new session (e.g., signature) keys that are valid only within that specific session.

The mechanism described above is quite natural as in the real world the parties are allowed to see signed messages passing on the channel and determine their actions based on these messages. The same should be allowed in the ideal world. Hence, we enhance the ideal functionality $\mathcal{F}$ by constructing $\mathcal{F}^\star$ which acts exactly as $\mathcal{F}$ as described above.

c) Collusion preservation with fallback security: Despite being simple and optimal in terms of round complexity, the protocol $\Pi^{\mathtt{HT}}$ above suffers from a big limitation. That is, if the hardware tokens are corrupted (e.g., the secret keys are leaked by the token manufacturer) then not only the CP-property is lost, but we cannot even guarantee to protect the honest parties' inputs. To rectify this issue we propose a protocol $\Pi^{\mathtt{HT\text{-}FBS}}$ that protects the input of the honest parties (in a standard GUC-security sense) even in the case where: 1) the adversary knows all the secret keys of the hardware tokens (including those held by honest parties) and 2) the malicious parties can arbitrarily modify or replace their own hardware tokens.

This protocol achieves a similar fallback security as the original collusion preserving protocol in the mediated model: When tokens are not compromised—and aborts are either excluded or deterred by means of incentives (see below)—then the protocol is collusion-preserving; and in any case (i.e., even when tokens are compromised) the protocol remains (G)UC secure—i.e., any profile of adversaries can be simulated by a monolithic simulator. We refer to a protocol that has such security guarantees as a *fallback secure* protocol.

Our fallback protocol guarantees also identifiable abort and unanimous abort for functionalities that guarantee termination, even when tokens can be broken. Interestingly, these properties are impossible to achieve in the analogous scenario of a corrupted mediator in the mediated model. To obtain such a protocol we use as a main building block a protocol Π^{MPC} that is secure against a malicious adversary and which enables identifiable (unanimous) abort. Each token computes protocol messages on behalf of its owner. In addition, the randomness used is jointly decided by the owner of the token and by the token itself.

More precisely, it is the XOR of the randomness produced by the hardware token, and a random string given at runtime by the token's owner. Intuitively, this means even if an honest party's token leaks its secret keys, it will still use an honestly-generated random string in the protocol. Thus, the party's input is protected by the security of the underlying MPC protocol, even if a token's secret keys are leaked. On the other hand, if all hardware tokens are uncompromised, then the randomness of any party in the MPC protocol becomes unknown and untamperable to everyone. Thus, no malicious party can send subliminal messages without being detected and causing an abort.

For sake of completeness, we also sketch a protocol which, although less round-efficient and without identifiable abort, preserves standard (GUC) security even when tokens are *fully* compromised—that is, not trusted, isolated, nor tamper-resilient. This is a stronger version of a "compromised" token since the adversary may read the contents or even change the behavior of honest parties' tokens. We make a simple alteration to our solution: Any computation performed by the token, will instead be done via a secure two-party protocol between the token and its owner. In more detail, each party p_j runs a secure 2-party protocol with his token, to obtain the next message of the protocol Π^{MPC}. When tokens are uncompromised, the behavior of this solution is the same as in our original solution, achieving CP. When tokens are compromised, fallback security follows from the security of 2-party protocol which ensures the tokens learn nothing. This solution, however, cannot achieve identifiable abort against fully-compromised tokens since any token can, e.g., be forced to stop any interactions. This also means the solution can only implement functionalities with abort, even with honest majority.

d) How to deal with aborting adversaries: We note that the above CP-protocols cannot prevent, for example, a corrupt poker player from simply sending an encrypted message "I have an ace of spades; let's collude!" over the broadcast channel. While the protocol could detect such invalid messages and abort, this attack already breaks CP and seem unavoidable if no assumptions are made on the network topology (e.g., the mediated model [1], [2], [4]) and on the honesty of the network nodes. In this work we circumvent the above issue by considering an *incentive-driven* (rational) attacker. That is, we define a security notion called CPAP (collusion preserving attack payoff secure) that captures the fact that some adversarial actions—in our case aborts—are not "for free" and instead incur a negative payoff. For example, as our protocols have (publicly) identifiable abort, a judge in the poker game example can identify and penalize an adversarial party causing the abort. Similar to the rational protocol design (RPD) framework [28], CPAP considers a CP-well-formed functionality $\mathcal{F}$, and a relaxed functionality $\langle\mathcal{F}\rangle$ that acts the same as $\mathcal{F}$ except it explicitly includes weaknesses that allow a simulator to collude or abort. Then, we define a value function v mapping the joint view of a set of simulators[8] interacting with the relaxed functionality $\langle\mathcal{F}\rangle$ and the environment $\mathcal{Z}$, to a real-valued *payoff*. Intuitively, the real utility gained/lost by a set of adversaries for a given protocol is the payoff maximized over all environments, and minimized over all sets of simulators that successfully emulate the adversaries in the environment. For our CP protocols, we show that collusion always causes the real world execution to abort (and identify a corrupt party). Thus, when the cost of abort exceeds the gains from colluding, a rational attacker will never collude (otherwise simulator can trigger an abort in $\langle\mathcal{F}\rangle$)—intuitively it means that our protocol implements $\mathcal{F}$ for rational attackers.

e) From "ideal" to "real" payoffs.: In Sec. VI-C we construct a penalization protocol Π_{pen} which runs the computation of our CP protocols $\Pi^{\mathtt{HT}}$ or $\Pi^{\mathtt{HT\text{-}FBS}}$, and in addition penalizes the adversary when he colludes (which can be done only by triggering an abort). For a natural class of utilities for the protocol designer and attacker, Π_{pen} enjoys both CPAP and CPIC, following similar arguments as in Sec. VI-B.

IV. Preliminaries

We denote the security parameter by λ and use "||" as concatenation operator (i.e., if a and b are two strings then

[8]Recall that in CP, adversaries are not monolithic, so we consider a set of adversaries/simulators instead of one single adversary/simulator.

by $a||b$ we denote the concatenation of a and b). For a finite set Q, $x \xleftarrow{\$} Q$ denotes a sampling of x from Q with uniform distribution. When it is necessary to refer to the randomness r used by and algorithm A we use the following notation: $A(\cdot; r)$. We denote with $[n]$ the set $\{1, \ldots, n\}$, with $\mathbb{F}$ an arbitrary (but fixed) finite field and with $\mathbb{N}$ the set of non-negative integers. We assume familiarity with the notions of secure-function evaluation, strong signatures, secret key encryption, negligible functions and pseudorandom functions and refer to the appendix for more detail.

1) Security with identifiable (unanimous) abort: In this work we consider the notion of secure function evaluation (SFE) with *identifiable abort*. This notion allows the computation to fail (abort), but ensures that when this happens all the honest parties are informed about it, and they also agree on the index i of some corrupted party $p_i \in \mathcal{P}$ [36]. We denote the ideal functionality for the evaluation of the function f that captures the property of identifiable abort with $\mathcal{F}_{\mathsf{IDA}}^{f}$ (where IDA stands for identifiable abort). We also consider the notion of *unanimous abort*. This guarantees that either all or none of the honest parties abort. We denote the ideal functionality for the evaluation of the function f that captures the property of unanimous abort with $\mathcal{F}_{\mathsf{UNA}}^{f}$ (where UNA stands for unanimous abort). We refer to the appendix for formal definitions of $\mathcal{F}_{\mathsf{IDA}}^{f}$ and $\mathcal{F}_{\mathsf{UNA}}^{f}$. We refer to [21] for a detailed discussion of the notions of security with non-unanimous (aka selective), unanimous, and identifiable abort, and their relation.

We will assume synchronous computation, i.e., our protocols proceed in rounds, where in each round: the uncorrupted parties generate their messages for the current round, as described in the protocol; then the messages addressed to the corrupted parties become known to the adversary; then the adversary generates the messages to be sent by the corrupted parties in this round; and finally, each uncorrupted party receives all the messages sent in this round. Although our treatment is in the (G)UC setting, to avoid overcomplicating the exposition, we will use the standard round-based language of [14], [48] to specify our protocol. Notwithstanding, such specifications can be directly translated to the synchronous UC model of Katz *et al.* [38] by assuming a clock functionality and bounded (zero) delay channels. We refer the interested reader to [38] for details.

2) Collusion-preserving computation: We now recall the notion of collusion-preserving computation proposed in [2], which we refer the reader to for a more thorough discussion on CP. For n the number of parties and $\mathcal{I} \subseteq [n]$, denote by $\mathcal{A}_{\mathcal{I}}$ the set of adversaries, i.e. ITMs $\{\mathcal{A}_i\}_{i \in \mathcal{I}}$, where $\mathcal{A}_i$ denotes the adversarial strategy of party p_i. In collusion-preserving computation, instead of one monolithic adversary/simulator, we consider a set of (independent) PPT adversaries and simulators. In more detail, we require the following:

- **Split adversaries/simulators:** Instead of a monolithic adversary/simulator we consider a set of n (independent) PPT adversaries $\mathcal{A}_{[n]} = \{\mathcal{A}_i, i \in [n]\}$, where $\mathcal{A}_i$ corresponds to the adversary associated with the player i (and can corrupt at most this party). We also ask that for each $\mathcal{A}_i \in \mathcal{A}_{[n]}$ there exists an (independent) simulator Sim_i, who only has access to $\mathcal{A}_i$.
- **Corrupted-Set Independence:** We also require that the simulators do not depend on each other. In other words the code of Sim_i is the same for any set of adversaries $\mathcal{A}_{[n]}$ and $\mathcal{B}_{[n]}$ as long as $\mathcal{A}_i = \mathcal{B}_i$.

Similar to the GUC framework (but in contrast to plain UC) we distinguish between two types of functionalities: resources, denoted with capital calligraphic font as in $\mathcal{R}$, and shared functionalities, denoted with an additional over-line as in $\bar{\mathcal{G}}$. Formally, a resource $\mathcal{R}$ maintains state only with respect to a single instance of a protocol, while a shared functionality $\bar{\mathcal{G}}$ can maintain state across protocol instances. For example, concurrent executions can maintain shared state via e.g., a global CRS or global PKI as long as they are modeled as shared functionalities. However, although concurrent instances of a protocol π may use the same resource $\mathcal{R}$, the behavior of $\mathcal{R}$ in one execution of π must be independent of all other executions of π (and more generally of all other concurrent protocols instantiated by the environment). For clarity, in the remainder of this work we will usually refer to shared functionalities simply as *setup*, and protocols that share state across executions only through some setup $\bar{\mathcal{G}}$ as $\bar{\mathcal{G}}$-*subroutine respecting*. We denote by $\text{CP-EXEC}^{\mathcal{R}}_{\Pi,\mathcal{A},\mathcal{Z}}$ the output of the environment $\mathcal{Z}$ in the execution of Π with adversaries $\mathcal{A} := \mathcal{A}_{[n]}$ in the $\mathcal{R}$-hybrid model. We say that a protocol Π is $\mathcal{R}$-*exclusive* if it makes use of no resources or shared functionality other than $\mathcal{R}$. We note that unlike (G)UC, CP limits parties to communicate with at most one single instance of the resource. Intuitively, if $\mathcal{F}$ is a one-bit channel, then the simulator only using one instance of $\mathcal{F}$ has a completely different meaning in terms of collusion-preservation to the simulator using unlimited calls to $\mathcal{F}$.

Definition 1 (Collusion Preservation [2])**.** *Let $\bar{\mathcal{G}}$ be a setup, $\mathcal{R}$ and $\mathcal{F}$ be n-party resources, Π be a $\{\bar{\mathcal{G}}, \mathcal{R}\}$-exclusive protocol and ϕ be a $\{\bar{\mathcal{G}}, \mathcal{F}\}$-exclusive protocol (both with n parties). Then we say that Π collusion-preservingly (CP) emulates ϕ in the $\{\bar{\mathcal{G}}, \mathcal{F}\}$-hybrid world, if there exists a collection of efficiently computable transformations $\mathsf{Sim} = \{\mathsf{Sim}_1, \ldots, \mathsf{Sim}_n\}$ mapping ITMs to ITMs such that for every set of adversaries $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, and every PPT environment $\mathcal{Z}$ the following holds:* $\text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi,\mathcal{A},\mathcal{Z}} \approx \text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\phi,\mathsf{Sim}(\mathcal{A}),\mathcal{Z}}$

Following [2], we distinguish between the notion of *emulation* and its special case *realization*. For a functionality $\mathcal{F}$ we denote by $\mathcal{D}_i^{\mathcal{F}}$ the ith dummy $\mathcal{F}$-hybrid protocol which simply acts as a transparent conduit between the ith honest and adversarial interfaces of $\mathcal{F}$ and the environment $\mathcal{Z}$. Specifically, $\mathcal{D}_i^{\mathcal{F}}$ forwards all messages it receives from $\mathcal{Z}$ to the $\mathcal{F}$ (where the choice of adversarial or honest interface is specified by $\mathcal{Z}$) and vice-versa. If for functionality $\mathcal{F}$, an $\mathcal{R}$-hybrid protocol π CP-emulates $\mathcal{D}_i^{\mathcal{F}}$ then we say that π *realizes* $\mathcal{F}$ (in the $\mathcal{R}$-hybrid world).

3) Rational protocol design: The goal of the rational protocol design framework (RPD) [6], [28] is to model security of protocols against incentive-driven attackers. In RPD, a protocol designer $\mathtt{D}$ engages in an *attack game* with an attacker $\mathtt{A}$. The designer first chooses a n-party protocol $\Pi \in \mathtt{ITM}^n$. Then based on Π, the attacker decides on an adversarial strategy $\mathcal{A}$ to attack Π. Each pair $(\Pi, \mathcal{A} = \mathtt{A}(\Pi))$ (called a *strategy profile*) induces a utility for the designer and the attacker. Consistent with [28], we consider an attack game $\mathcal{G}_{\mathcal{M}}$, where $\mathcal{M}$ is the attack model $\mathcal{M} = (\mathcal{F}, \langle\mathcal{F}\rangle, v_{\mathtt{A}})$, a vector of parameters of the game. $\mathcal{F}$ is the functionality which the designer would like to achieve, and $\langle\mathcal{F}\rangle$ is a relaxed version of $\mathcal{F}$ in the sense that it includes extra commands that break certain security properties of $\mathcal{F}$. The value function $v_{\mathtt{A}}$ allows us to define utilities of the attacker. It assigns payoffs when certain events occur in the ideal world, such as when the simulator uses the extra commands in $\langle\mathcal{F}\rangle$ to help him complete a successful simulation. Now, our goal is ensure that it is not in the attacker's best interest to force the simulator to use weaknesses of $\langle\mathcal{F}\rangle$. This is captured by the notion of *attack-payoff security*. A protocol Π is *attack-payoff secure* if no adversarial strategy attacking Π can achieve more utility than an adversary attacking the "dummy" protocol that just uses the functionality $\mathcal{F}$ (i.e. without weaknesses of $\langle\mathcal{F}\rangle$). In other words, Π is "as secure as" the dummy (trivially secure) protocol in this model. As in [6], we can augment $\mathcal{M}$ with the designer's value function $v_{\mathtt{D}}$, and consider whether the protocol is *incentive compatible* for the designer.

V. CP MPC WITH NON-ABORTING ADVERSARIES

In this section we present our protocol $\Pi^{\mathtt{HT}}$ for any CP-well-formed functionality under the following assumptions: 1) each party has access to a hardware token (which we describe as one global ideal functionality in Fig. 1) 2) all communication is done over authenticated broadcast, and 3) adversarial parties do not make the protocol abort. For simplicity we restrict ourselves to non-reactive functionalities, also known as secure function evaluation. (The general case can be reduced to this case using a suitable form of secret sharing to maintain the secret intermediate states of the reactive functionality.) Moreover, we describe all our protocols in a round based, synchronous manner, where messages sent in some round are delivered by the beginning of the next round. We first introduce some additional notation:

- $\mathsf{sid} \in \mathbb{N}$ uniquely identifies an execution of $\Pi^{\mathtt{HT}}$.
- Set of parties $\mathcal{P} = \{p_1, \ldots, p_n\}$ running $\Pi^{\mathtt{HT}}$ compute the function f.
- We call *leader* the party that is in charge to run a special code and we assume w.l.o.g. that the leader is $p_n \in \mathcal{P}$.
- $T^{\mathtt{HT}}$ denotes the global token functionality.

For our construction we use the following tools: Pseudorandom functions: $\mathsf{PRF}_0 : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^\lambda$, $\mathsf{PRF}_1 : \{0,1\}^\lambda \times \{0,1\}^{2\lambda} \to \{0,1\}^{4\lambda}$, $\mathsf{PRF}_2 : \{0,1\}^\lambda \times \{0,1\}^{2\lambda} \to \{0,1\}^{(n+2)\lambda}$, $\mathsf{PRF}_3 : \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^{(4n-2)\lambda}$; A strong unforgeable signature scheme $\Sigma = (\mathsf{Kgen}, \mathsf{Sign}, \mathsf{Ver})$; A secret-key encryption scheme $\Pi^{\mathsf{SK}} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$.

The global setup $\bar{\mathcal{G}}$ is represented by the token functionality $T^{\mathtt{HT}}$. We note that we use one functionality to *emulate* the behavior of the hardware tokens held by the parties that run the protocol. This token functionality replies to each party $p_i \in \mathcal{P}$ using the appropriate code and keys depending to the identity of the calling party (i.e. the functionality discriminates between leader and non-leader parties). To not overburden the notation, in the formal construction we denote the identity of a party $p_i \in \mathcal{P}$ with i. Moreover, the token functionality exports as public information the master public key mpk, and keep as part of its secret state the master secret key msk together with with the PRF keys K_0, K_1, K_2, K_3. The parties are allowed to communicate only via a broadcast channel denoted by $\mathcal{B}$ (c.f. App. H for formal definitions). We provide a formal description of $T^{\mathtt{HT}}$ in Fig. 1. The complete formal description of the the protocol $\Pi^{\mathtt{HT}}$ for the non-leader party is proposed in Fig. 2, and the protocol run by the leader parties is provided in Fig. 3. We assume that the ideal functionality $\mathcal{F}$ that we wish to realize is registered to the token functionality. In addition, upon receiving the command $(\mathsf{GetTrapdoor}, \mathsf{sid})$, $\mathcal{F}$ sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to the token functionality if sid is equal its session id, and forwards the answer to the ideal adversary. We recall that this trapdoor allows us to capture the broadcast channel, on which all parties see the exact same signed messages. Equipping the ideal functionality with the trapdoor command translates this *real-world leakage* in the ideal world. We also recall that the functionality leaks to the simulators messages that are valid within one specific session without harming the token functionality globally. We now prove that $\Pi^{\mathtt{HT}}$ is collusion-preserving against non-aborting adversaries for well-formed functionalities. Formally, we prove the following:

Theorem 1. *Let $\bar{\mathcal{G}} = T^{\mathtt{HT}}$ be the setup as defined above, $\mathcal{R} = \mathcal{B}$ (broadcast) and $\mathcal{F}$ be n-party resources where $\mathcal{F}$ is a CP-well-formed functionality. Then the $\{\bar{\mathcal{G}}, \mathcal{R}\}$-exclusive protocol $\Pi^{\mathtt{HT}}$ (described by Fig. 2 and 3) CP realizes $\mathcal{F}$ in the $\mathcal{R}$-hybrid world assuming non-aborting adversaries.*

Proof sketch. We refer to App. B for the full proof. For simplicity we assume that only the leader is honest. To prove this theorem, we need to show a collection of efficiently computable transformations $\mathsf{Sim} = \{\mathsf{Sim}_1, \ldots, \mathsf{Sim}_n\}$ that satisfy Definition 1. For $i = 1, \ldots n$, the simulator $\mathcal{S}_i = \mathsf{Sim}(\mathcal{A}_i)$ queries $(\mathsf{GetTrapdoor}, \mathsf{sid})$ $\mathcal{F}^\star$ with command $(\mathsf{GetTrapdoor}, \mathsf{sid})$. $\mathcal{F}^\star$ checks that sid is equal to its session id. If so, then $\mathcal{F}^\star$ sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to $T^{\mathtt{HT}}$. $T^{\mathtt{HT}}$ then generates a set of encryptions of 0^λ, a set of signing/verification keys and uses them to authenticate these encryptions (see the bottom of Fig. 1). We note that each simulator will obtain the same set of ciphertexts and verification keys. This is crucial for the proof to go through, as each individual simulator

The token functionality is parameterized by a set of parties $\mathcal{P}$ and by a list $\overline{\mathcal{F}}$ of ideal functionality programs. The functionality manages the keys $(\mathsf{mpk}, \mathsf{msk})$ for the signature scheme Σ and the PRF keys K_0, K_1, K_2, K_3.

If $I = (\mathsf{Get\_key}, \mathsf{sid})$ is received return mpk to the caller.

Input phase for non-leader parties.

If $I = (\mathsf{Input}, \mathsf{sid}, x, f)$ is received from a non-leader party p_j then do the following.

- If $\mathsf{ctr}_j^{\mathsf{sid}}$ is not defined then define it and set $\mathsf{ctr}_j^{\mathsf{sid}} \leftarrow 1$ otherwise output $\perp$ and stop.
- Compute $R_0 \leftarrow \mathsf{PRF}_0(K_0, \mathsf{sid})$ and $\mathsf{Kenc}^{\mathsf{sid}} \leftarrow \mathsf{Gen}(1^\lambda; R_0)$
- Compute $R_1 \leftarrow \mathsf{PRF}_1(K_1, \mathsf{sid}||j)$ and parse R_1 as 4 strings of λ bits each $\mathsf{rs}^1||\mathsf{rs}^2||r^1||r^2$.
- $(\mathsf{sigk}_j^{\mathsf{sid}}, \mathsf{vk}_j^{\mathsf{sid}}) \leftarrow \mathsf{Kgen}(1^\lambda; \mathsf{rs}^1)$
- $\mathsf{cert}_j \leftarrow \mathsf{Sign}(\mathsf{msk}, \mathsf{vk}_j^{\mathsf{sid}}||\mathsf{sid}||j; \mathsf{rs}^2)$
- Compute $\overline{x} \leftarrow \mathsf{Enc}(\mathsf{Kenc}^{\mathsf{sid}}, x; r^1)$, $\sigma \leftarrow \mathsf{Sign}(\mathsf{sigk}_j^{\mathsf{sid}}, \overline{x}||f; r^2)$ and output $(\overline{x}, f, \mathsf{vk}_j, \sigma, \mathsf{cert}_j)$.

Input/output phase for the leader party.

If $I = (\mathsf{Input}, \mathsf{sid}, x_n, f, \mathsf{sid}, (\overline{x}_1, \mathsf{vk}_1, \sigma_1, \mathsf{cert}_1), \ldots, (\overline{x}_{n-1}, \mathsf{vk}_{n-1}, \sigma_{n-1}, \mathsf{cert}_{n-1}))$ is received from the leader party p_n then check if for all $j \in [n-1]$ $\mathsf{Ver}(\mathsf{vk}_j, \overline{x}_j||f, \sigma_j) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk}_j||\mathsf{sid}||j, \mathsf{cert}_j) = 1$. If it is not, then output $\perp$ and stop, otherwise act as follows.

- If $\mathsf{ctr}_n^{\mathsf{sid}}$ is not defined then define it and set $\mathsf{ctr}_n^{\mathsf{sid}} \leftarrow 1$ else output $\perp$ and stop.
- $R_2 \leftarrow \mathsf{PRF}_2(K_2, \mathsf{sid}||n)$ and parse R_2 as $n+2$ strings of λ bits each $\mathsf{rs}^1||\mathsf{rs}^2||r_1||r_2||\ldots||r_{n-1}||r^\star$.
- $(\mathsf{sigk}_n^{\mathsf{sid}}, \mathsf{vk}_n^{\mathsf{sid}}) \leftarrow \mathsf{Kgen}(1^\lambda; \mathsf{rs}^1)$
- $\mathsf{cert}_n \leftarrow \mathsf{Sign}(\mathsf{msk}, \mathsf{vk}_n^{\mathsf{sid}}||\mathsf{sid}||n; \mathsf{rs}^2)$
- Compute $R_0 \leftarrow \mathsf{PRF}_0(K_0, \mathsf{sid})$ and $\mathsf{Kenc}^{\mathsf{sid}} \leftarrow \mathsf{Gen}(1^\lambda; R_0)$.
- For $j = 1, \ldots, n-1$ compute $x_j \leftarrow \mathsf{Dec}(\mathsf{Kenc}^{\mathsf{sid}}, \overline{x}_j)$.
- Compute $y_1, \ldots, y_n \leftarrow f(x_1, \ldots, x_n)$.
- For $j = 1, \ldots, n-1$ compute $\overline{y}_j \leftarrow \mathsf{Enc}(\mathsf{Kenc}^{\mathsf{sid}}, y_j; r_j)$.
- $\sigma \leftarrow \mathsf{Sign}(\mathsf{sigk}_n^{\mathsf{sid}}, \overline{y}_1||\ldots||\overline{y}_{n-1}||f; r^\star)$;
- Output $(\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n), y_n$

Output phase for non-leader parties. If $I = (\mathsf{Output}, \mathsf{sid}, z)$ is received, parse z as $(\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ and do the following. If $\mathsf{Ver}(\mathsf{vk}_n, \overline{y}_1||\ldots||\overline{y}_{n-1}||f, \sigma) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk}_j||\mathsf{sid}||j, \mathsf{cert}_j) = 1$ then compute and output $\mathsf{Dec}(\mathsf{Kenc}^{\mathsf{sid}}, \overline{y}_j)$, output $\perp$ otherwise.

Trapdoor. If $I = (\mathsf{Trapdoor}, \mathsf{sid})$ is received from an instance of an ideal functionality in the list $\overline{\mathcal{F}}$ then do the following.

- If $\mathsf{ctr}_j^{\mathsf{sid}}$ is not defined then define it and set $\mathsf{ctr}_j^{\mathsf{sid}} \leftarrow 1$ otherwise output $\perp$ and stop.
- Pick $r_1||\ldots||r_{n-1}||r'_1||\ldots||r'_{n-1}||\mathsf{rs}_1^1||\mathsf{rs}_1^2||\ldots||\mathsf{rs}_n^1||\mathsf{rs}_n^2 \leftarrow \mathsf{PRF}_3(K_3, \mathsf{sid})$.
- For each $i \in [n]$ $(\mathsf{sigk}_i, \mathsf{vk}_i) \leftarrow \mathsf{Kgen}(1^\lambda; \mathsf{rs}_i^1)$, $\mathsf{cert}_i \leftarrow \mathsf{Sign}(\mathsf{msk}, \mathsf{vk}_i||\mathsf{sid}||i; \mathsf{rs}_i^2)$. For each $j \in [n-1]$ $e_j \leftarrow \mathsf{Enc}(\mathsf{pk}_n, 0^\lambda; r_j)$, $\sigma_j \leftarrow \mathsf{Sign}(\mathsf{sigk}_j, e_j; r'_j)$
- For each $j \in [n-1]$ compute $\overline{y}_j \leftarrow \mathsf{Enc}(\mathsf{pk}_j, 0^\lambda; r'_j)$
- $\sigma_n \leftarrow \mathsf{Sign}(\mathsf{sigk}_n, \overline{y}_1||\ldots||\overline{y}_{n-1}||f||\mathsf{sid})$.
- Return to the calling instance $\{\mathsf{vk}_i, \mathsf{cert}_i, \sigma_i\}_{i\in[n]}, \{e_i, \overline{y}_i\}_{i\in[n-1]}$.

Fig. 1. The functionality $T^{\mathtt{HT}}$ models the behaviour of the hardware tokens.

We assume that the party p_j is registered to the token functionality $T^{\mathtt{HT}}$ and that it obtains mpk by querying it with $I = (\mathsf{Get\_key}, \mathsf{sid})$. Each party is aware of the function that will be computed f, of the identifier of each execution sid, and of the parties involved in each of those executions $\mathcal{P}$.

Input.

- The party p_j on input $(\mathsf{Compute}, \mathsf{sid}, x)$ sends $(\mathsf{Input}, \mathsf{sid}, x, f)$ to $T^{\mathtt{HT}}$.
- Upon receiving the answer X from $T^{\mathtt{HT}}$, if $X = \perp$ then p_j outputs $\perp$ and stops. Otherwise, p_j sends X to p_n.

Output. The party p_j, upon receiving $z = (\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ from p_n sends $(\mathsf{Output}, \mathsf{sid}, z)$ to $T^{\mathtt{HT}}$. Upon receiving y from $T^{\mathtt{HT}}$, p_j outputs y.

Check-channel. The party p_j inspects all messages that are sent on the channel. If a message $(m, f', \mathsf{vk}, \sigma, \mathsf{cert})$ is received from a party p_i check if $f = f'$ and $\mathsf{Ver}(\mathsf{vk}_i, m||f, \sigma) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk}||\mathsf{sid}||i, \mathsf{cert}) = 1$. If it is not, then output $(\perp, p_i)$ and stop.

Fig. 2. Protocol executed by the party p_j.

Input/output

- The party p_n on input $(\mathsf{Compute}, \mathsf{sid}, x_n)$ collects messages from $p_{j\in[n-1]}$ and sends $I = (\mathsf{Input}, x_n, f, \mathsf{sid}, (\overline{x}_1, \mathsf{vk}_1, \sigma_1, \mathsf{cert}_1), \ldots, (\overline{x}_{n-1}, \mathsf{vk}_{n-1}, \sigma_{n-1}, \mathsf{cert}_{n-1}))$ to $T^{\mathtt{HT}}$.
- Upon receiving the answer Y from $T^{\mathtt{HT}}$, if $Y = \perp$ then output $\perp$ and stop. Otherwise parse Y as $((m, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n), y)$, send $(m, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ to all the parties in $\mathcal{P}$ and output y.

Check-channel. The party p_j inspects all messages that are sent on the channel. If a message $(m, f', \mathsf{vk}, \sigma, \mathsf{cert})$ is received from a party p_i check if $f = f'$ and $\mathsf{Ver}(\mathsf{vk}_j, m||f, \sigma) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk}||\mathsf{sid}||j, \mathsf{cert}) = 1$. If it is not, then output $(\perp, p_i)$ and stop.

Fig. 3. Protocol executed by the leader party p_n.

$\mathcal{S}_i$ internally runs the corrupted party p_i, and it will act on the behalf of the other $n-1$ parties using the authenticated cipthertexts received by $\mathcal{F}^\star$. $\mathcal{S}_i$ will also intercept all the queries the party p_i makes to $T^{\mathtt{HT}}$. Assuming that p_i is non-aborting, then at some point they will query $T^{\mathtt{HT}}$ with their input x_i. $\mathcal{S}_i$ now has the input of the corrupted party, and he will use it to query the ideal functionality. In addition, $\mathcal{S}_i$ sends to p_i (acting on the behalf of $T^{\mathtt{HT}}$) an encryption of 0 authenticated with respect to the verification key vk_i. Assuming that no party aborts, the simulation is successful since the messages generated by the individual simulator on the locally simulated broadcast channels are exactly the same for all the parties (unless the adversary breaks the signature scheme).

1) (Publicly) Identifiable abort: Another interesting property enjoyed by $\Pi^{\mathtt{HT}}$ is *identifiable abort*. A protocol run by a set of parties $\mathcal{P}$ is said to be secure with identifiable abort if it either computes according to its specification, or it aborts with the index of some corrupted party $p_i \in \mathcal{P}$ —i.e., every honest party learns the identity of a corrupted p_i. In $\Pi^{\mathtt{HT}}$ the adversary can only deviate from the protocol specification by

either sending a message authenticated with respect to a sid' or f' not equal to the correct sid or f the honest parties use, sending a message with an invalid signature or certificate, or fail to send a message. Each event is verifiable by honest parties, and even third parties not involved in the protocol. Indeed, with the master public key mpk, sid and function f, it is possible to claim who did abort in a run of $\Pi^{\mathtt{HT}}$ by just inspecting its transcript. Formally, the protocol $\Pi^{\mathtt{HT}}$ securely realizes the function $\mathcal{F}_{\mathsf{IDA}}^{f}$, where $\mathcal{F}_{\mathsf{IDA}}^{f}$ involves n parties. More interestingly, we can modify $\Pi^{\mathtt{HT}}$ to support an additional party p_{n+1} which takes no input, does not send any message and outputs a default value (e.g., 0). Since p_{n+1} knows the master public key mpk, she can check the validity of the signature and the certificate. Hence, she is able to identify an invalid message (in the case p_{n+1} is honest). That is, our protocol allows an *observer* of the protocol execution to identify a misbehaving party. Following [39] we refer to this property as *publicly identifiable abort*, and to p_{n+1} as a *judge*. The code of the judge can be used by anyone who has the public setup and wants to follow the protocol execution and decide who aborted the protocol given the parties' messages.

2) Note on implementing $T^{\mathtt{HT}}$ with real hardware tokens: Following the approach of [3], we describe the behavior of the hardware tokens (HTs) by means of a single ideal functionality with a single set of shared keys. In particular, the master signing/verification keys allow tokens and parties to verify whether a message was signed by a hardware token. In practice, such functionality can be replaced by tokens that each has its own secret information (e.g., the HTs do not need share the same master signing key). More precisely, only a global master verification key needs to be shared among the tokens. Each hardware token manages its own signing and verification keys, msk_i and mpk_i, along with a certificate c_i which is a signature of mpk_i created by the HT manufacturer's global master secret key. An example of this approach is Intel SGX processors, where each processor has a unique attestation key and "endorsement certificate" from the manufacturer [22]. To authenticate a message, HT HT_i signs it with its msk_i, and sends the signature together with his master verification key vk_i and the certificate c_i. Anybody that has the global master verification key can verify that the certificate c_i is valid for vk_i, and that the signature issued by HT_i is valid w.r.t. vk_i.

Another important part of $T^{\mathtt{HT}}$ is to ensure that the session ID sid must not be reused in different protocol sessions. The most simple solution stores all sids that the token has seen. Thus, if the token cannot store more sids, it will not be able to verify the freshness of new sids and must stop responding all together. This in effect makes the token no longer usable for our CP protocol. We present two alterations to improve upon the space usage of the simple solution. First, the token can transfer the burden of storing the sids to an external memory, in a hash chain data structure [43]. The token stores the head and tail of the hash chain, which ensures that no malicious party can tamper with the sids on the external memory. The solution only requires a small (i.e., constant in the number of sids) amount of storage. However, verifying a sid requires interaction with the external memory to retrieve the hash chain—to use minimal space, the token may choose to retrieve the chain one hash at a time. To eliminate interactions with external memory, one may also opt for a Bloom filter [13]. This solution trades off the need for interaction with the possibility (depending on the space allocated to the filter) of falsely marking some sids as "used". This however does not impact security, as it is equivalent to the token having seen a session with this sid.

A. Fallback solution when tokens can be compromised

While simple and optimal in terms of round complexity, the protocol $\Pi^{\mathtt{HT}}$ cannot guarantee any form of security when the hardware tokens are corrupted/compromised. In this section we propose a CP protocol $\Pi^{\text{HT-FBS}}$ that provides fallback security (in a standard GUC-security sense) in the following corruption model: 1) the secret keys of the tokens, including those of honest parties, can be leaked and 2) the memory/code of corrupt parties' tokens can be read/modified completely. Let $\mathcal{F}$ be the function that the parties wish to compute. For our construction we use the following tools.

- Pseudo-random functions $\mathsf{PRF}_0 : \{0,1\}^{\lambda} \times \{0,1\}^{2\lambda} \to \{0,1\}^{(2m+4)\lambda}$ and $\mathsf{PRF}_1 : \{0,1\}^{\lambda} \times \{0,1\}^{\lambda} \to \{0,1\}^{((m+3)n+1)\lambda}$
- A strong unforgeable signature scheme $\Sigma = (\mathsf{Kgen}, \mathsf{Sign}, \mathsf{Ver})$.
- A n-party MPC protocol $\Pi^{\mathsf{MPC}} = (\mathsf{Next}_1, \ldots, \mathsf{Next}_n)$ that GUC-realizes functionality $\mathcal{F}$.

The global setup $\bar{\mathcal{G}}$ is represented by the token functionality $T^{\text{HT-FBS}}$. Similar to the token functionality of the previous section, $T^{\text{HT-FBS}}$ has a master public key mpk and a secret state consisting of the corresponding master secret key msk and the PRF keys K_0, K_1. We assume without loss of generality that the setup required to run Π^{MPC} is part of $T^{\text{HT-FBS}}$. We denote our protocol with $\Pi^{\text{HT-FBS}}$ (Fig. 5) and provide a formal description of $T^{\text{HT-FBS}}$ in Fig. 4.

Security of $\Pi^{\text{HT-FBS}}$:

We summarize the properties of the protocol $\Pi^{\text{HT-FBS}}$.

1) If the hardware tokens are not compromised, and no party aborts, then $\Pi^{\text{HT-FBS}}$ is collusion-preserving.
2) If the hardware tokens are compromised and Π^{MPC} GUC realizes $\mathcal{F}_{\mathsf{AB}}^{f}$ with $\mathsf{AB} \in \{\mathsf{IDA}, \mathsf{UNA}\}$, then $\Pi^{\text{HT-FBS}}$ GUC realizes $\mathcal{F}_{\mathsf{AB}}^{f}$
3) If the hardware tokens are not compromised (but the malicious parties may abort), then $\Pi^{\text{HT-FBS}}$ GUC realizes the functionality $\mathcal{F}$ with publicly identifiable abort.

The properties 1 and 2 above enable the fallback security of $\Pi^{\text{HT-FBS}}$. In addition, the second property states that in the case of corrupted tokens, $\Pi^{\text{HT-FBS}}$ inherits all the properties of Π^{MPC} (e.g., identifiable abort). We note that if the MPC protocol guarantees fairness (or even output delivery), this property would be held by $\Pi^{\text{HT-FBS}}$ as well. The third property states that if an adversarial party aborts then the CP property might

be lost, but the input of the honest parties are protected. We capture the case where the hardware tokens are compromised by considering the token functionality $\overline{T}^{\text{HT-FBS}}$ instead of $T^{\text{HT-FBS}}$. $\overline{T}^{\text{HT-FBS}}$ extends $T^{\text{HT-FBS}}$ with the additional command Tamper. If the adversary queries the token functionality with Tamper then $\overline{T}^{\text{HT-FBS}}$ leaks to the adversary its secret state (i.e., the master secret key msk and the PRF keys). Given the master secret key, the adversary can authenticate any message he wants and therefore acts on the behalf of the hardware token. To formally prove that $\Pi^{\text{HT-FBS}}$ is fallback secure we need to prove the following two lemmata.

Lemma 1. *Let $\bar{\mathcal{G}} = T^{\text{HT-FBS}}$ be the setup, $\mathcal{R} = \mathcal{B}$ (broadcast) and $\mathcal{F}$ be n-party resources where $\mathcal{F}$ is a CP-well-formed functionality. Then the $\{\bar{\mathcal{G}}, \mathcal{R}\}$-exclusive protocol $\Pi^{\text{HT-FBS}}$ (described by Fig. 5) CP realizes $\mathcal{F}$ in the $\mathcal{R}$-hybrid world assuming that no parties abort.*

Lemma 2. *Let Π^{MPC} be a protocol that GUC-realizes the n-party functionality $\mathcal{F}^{\mathsf{AB}}$ with $\mathsf{AB} \in \{\mathsf{IDA}, \mathsf{UNA}\}$ that exclusively uses $\mathcal{B}$ as a resource. Let $\bar{\mathcal{G}} = \overline{T}^{\text{HT-FBS}}$ and $\mathcal{R} = \mathcal{B}$ then $\forall \mathcal{A}\ \exists \mathsf{Sim}\ \forall \mathcal{Z}\ \text{EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi^{\text{HT-FBS}},\mathcal{A},\mathcal{Z}} \approx \text{EXEC}^{\bar{\mathcal{G}},\mathcal{F}^{\mathsf{AB}}}_{\mathsf{Sim},\mathcal{Z}}$*

Theorem 2. *Let $\bar{\mathcal{G}} = T^{\text{HT-FBS}}$ be the setup, $\mathcal{R} = \mathcal{B}$ and $\mathcal{F}$ be n-party resources where $\mathcal{F}$ is a CP-well-formed functionality. Then, the $\{\bar{\mathcal{G}}, \mathcal{R}\}$-exclusive protocol $\Pi^{\text{HT-FBS}}$ (described by Fig. 5) GUC realizes $\mathcal{F}^{\mathsf{IDA}}$.*

We refer to App. B for the formal proof of the above lemmata and theorem.

1) Remark on fully-compromised tokens: As discussed in the introduction, to achieve fallback security against fully-compromised (that is, untrusted, non-tamper-resilient, non-isolated) tokens, any computation performed by the token must instead be done through a 2-party protocol between the token and the party holding it. To be specific, this 2-party protocol will implement the following functionality: On request to compute: (1) if round $\ell = 0$: Take input x_j from party p_j and $K_0, \mathsf{mpk}, \mathsf{msk}$ from p_j's token. It stores $x_j, K_0, \mathsf{mpk}, \mathsf{msk}$. (2) If $\ell \leq m$: Take input the messages from the current round from p_j, if a message fails to authenticate then the token stops and p_j aborts. (3) Finally, output the next message $(\mathtt{msg}^{\ell}_j, \cdots)$ as the token would, to p_j.

2) Remark on malicious broadcast: We also give intuition on what happens if our authenticated broadcast resource is malicious. First, we will invariably lose CP, since a malicious broadcast resource can provide a private communication channel for corrupt parties, trivially allowing collusion. However, like with a malicious mediator in [2], we still achieve (G)UC security without unanimous abort/fairness. This is an interesting feature of our model and results: our security degrades more gradually with respect to the underlying assumptions than with a corrupted mediator. When only the tokens are compromised but authenticated broadcast is honest, then we have (G)UC security with identifiable abort—impossible with a corrupt mediator. If broadcast is malicious then we lose identifiable/unanimous abort and fairness—e.g., the malicious broadcast can pass an unauthenticated invalid message to just one honest party, making only this party abort, and it can also refuse to pass the output to honest parties. Intuitively, security is preserved even with malicious broadcast since messages

The functionality manages the keys $(\mathsf{mpk}, \mathsf{msk})$ for the signature scheme Σ. The functionality manages also the PRF keys K_0, K_1.

If $I = (\mathsf{Get\_key}, \mathsf{sid})$ is received return mpk to the caller.

Input phase. If $I = (\mathsf{Input}, \mathsf{sid}, x_j, R_j)$ is received from some party p_j, then do the following,

- If $\mathsf{ctr}^{\mathsf{sid}}_j$ is not defined, then define it and $\mathsf{ctr}^{\mathsf{sid}}_j \leftarrow 1$, otherwise output $\perp$ and stop.
- Compute $R^{\mathsf{sid}}_0 \leftarrow \mathsf{PRF}_0(K_0, \mathsf{sid}||j) \oplus R_j$ and parse R^{sid}_0 as $(2m+4)$ strings of λ bits $\mathsf{rs}^1_j||\mathsf{rs}^2_j||\rho^1_j||\ldots||\rho^{m+1}_j||r^1_j||\ldots||r^{m+1}_j$.
- $(\mathsf{sigk}^{\mathsf{sid}}_j, \mathsf{vk}^{\mathsf{sid}}_j) \leftarrow \mathsf{Kgen}(1^{\lambda}; \mathsf{rs}^1_j)$
- $\mathsf{cert}^{\mathsf{sid}}_j \leftarrow \mathsf{Sign}(\mathsf{msk}, \mathsf{vk}_j||\mathsf{sid}||j; \mathsf{rs}^2_j)$
- Output the first message $(\mathtt{msg}^1_j, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma, \mathsf{vk}_j, \mathsf{cert}_j)$, where $\mathtt{msg}^1_j = \mathsf{Next}_j(1^{\lambda}, x_j, \rho^1_j, \perp)$ and $\sigma \leftarrow \mathsf{Sign}(\mathsf{sigk}_j, \mathtt{msg}^1_j||\Pi^{\mathsf{MPC}}||\ell; r^1_j)$

Next message function.

If $I = (\mathsf{NextMsg}, \mathsf{sid}, \{\mathtt{msg}^{\ell}_i, \sigma^{\ell}_i, \mathsf{vk}_i, \mathsf{cert}_i\}_{i\in[n]})$ is received from p_j then do the following.

- If $\ell > m$ or $\mathsf{ctrap}^{\mathsf{sid}}_j = 1$ then output $\perp$ and stop, else continue with the following steps.
- Parse R^{sid}_0 as $\mathsf{rs}^1_j||\mathsf{rs}^2_j||\rho^1_j||\ldots||\rho^{m+1}_j||r^1_j||\ldots||r^{m+1}_j$.
- Store $\mathtt{msg}^{\ell}_{\mathsf{sid}} = \{\mathtt{msg}^{\ell}_i\}_{i\in[n]}$.
- For all $i \in [n]$ check if $\mathsf{Ver}(\mathsf{vk}_i, \mathtt{msg}^{\ell}_i||\Pi^{\mathsf{MPC}}||\ell, \sigma^{\ell}_i) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk}_i||\mathsf{sid}||i, \mathsf{cert}_i) = 1$. If it is not then output $(p_i, \perp)$ and stop. Otherwise continue with the following steps.
- Set $\ell \leftarrow \ell + 1$, compute $\mathtt{msg}^{\ell}_j = \mathsf{Next}_j(1^{\lambda}, x_j, \rho^{\ell}_j, \overline{\mathtt{msg}}^{<\ell})$ with $\overline{\mathtt{msg}}^{<\ell} = \{\mathtt{msg}^{\ell'}_i\}_{i\in[n],\ell'<\ell}$ where $\mathtt{msg}^{\ell'}_i$ is the message from p_i at round ℓ' that was stored in $\mathtt{msg}^{\ell'}_{\mathsf{sid}}$
- If $\ell = m+1$ then output $y_j \leftarrow \mathtt{msg}^{\ell}_j$
  else, output $(\mathtt{msg}^{\ell}_j, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma^{\ell}_j, \mathsf{vk}_j, \mathsf{cert}_j)$, where $\sigma^{\ell}_j \leftarrow \mathsf{Sign}(\mathsf{sigk}^{\mathsf{sid}}_j, \mathtt{msg}^{\ell}_j||\Pi^{\mathsf{MPC}}||\ell; r^{\ell}_j)$.

Trapdoor. If $I = (\mathsf{Trapdoor}, \mathsf{sid})$ is received from an instance of an ideal functionality in the list $\overline{\mathcal{F}}$ then do the following.

- If $\mathsf{ctr}^{\mathsf{sid}}_j$ is not defined, then define it, define $\mathsf{ctrap}^{\mathsf{sid}}_j$, set $\mathsf{ctr}^{\mathsf{sid}}_j \leftarrow 1$ and $\mathsf{ctrap}^{\mathsf{sid}}_j \leftarrow 1$. Else output $\perp$ and stop.
- Compute $\rho||r^1_1||\ldots||r^{m+1}_1||\ldots||r^1_n||\ldots||r^{m+1}_n||\mathsf{rs}^1_1||\mathsf{rs}^2_1||\ldots||\mathsf{rs}^1_n||\mathsf{rs}^2_n \leftarrow \mathsf{PRF}_1(K_1, \mathsf{sid})$.
- For all $i \in [n]$ $(\mathsf{sigk}_i, \mathsf{vk}_i) \leftarrow \mathsf{Kgen}(1^{\lambda}; \mathsf{rs}^1_i)$, $\mathsf{cert}_i \leftarrow \mathsf{Sign}(\mathsf{msk}, \mathsf{vk}_i||\mathsf{sid}||i; \mathsf{rs}^2_i)$.
- Use ρ to run the simulator of the MPC, S for the case where there are no corrupted party.
- Let $\{\mathtt{msg}^{\ell}_j\}_{j\in[n],\ell\in[m]}$ be the messages contained in the transcript obtained by the MPC simulator S. For all $j \in [n]$ and $\ell \in [m]$ computes $\sigma^{\ell}_j \leftarrow \mathsf{Sign}(\mathsf{sigk}_j, \mathtt{msg}^{\ell}_j||\Pi^{\mathsf{MPC}}||\ell; r^{\ell}_j)$.
- Return $\{\mathsf{vk}_i, \mathsf{cert}_i\}_{i\in[n]}, \{\mathtt{msg}^{\ell}_j, \sigma^{\ell}_j\}_{j\in[n],\ell\in[m]}$

Fig. 4. The functionality $T^{\text{HT-FBS}}$ models the behaviour of the hardware tokens, in our fallback secure protocol $\Pi^{\text{HT-FBS}}$.

We assume that the party p_j is registered to the global token functionality $T^{\text{HT-FBS}}$ and obtains mpk by querying it with $I = (\mathsf{Get\_key}, \mathsf{sid})$.

Input and next message generation.

- The party p_j on input $(\mathsf{Compute}, \mathsf{sid}, x)$ samples uniform random $R_j \in \{0,1\}^{(2m+4)\lambda}$ and sends $I = (\mathsf{Input}, \mathsf{sid}, x_j, R_j, \Pi^{\mathsf{MPC}})$ to $T_j^{\text{HT-FBS}}$.
- For each $\ell \in \{1, \ldots, m\}$:
  - Upon receiving message X from $T_j^{\text{HT-FBS}}$ check if $X = (\bot, p_{i'})$. If it is then output $(\bot, p_{i'})$ and stop, otherwise send X over broadcast.
  - Collect message $(\mathsf{msg}_i^\ell, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma_i^\ell, \mathsf{vk}_i, \mathsf{cert}_i)$ for round ℓ from each party $p_i \in [n] \backslash \{j\}$ and send $(\mathsf{NextMsg}, \mathsf{sid}, \{\mathsf{msg}_i^\ell, \sigma_i^\ell, \mathsf{vk}_i, \mathsf{cert}_i\}_{i \in [n]})$ to $T_j^{\text{HT-FBS}}$.

Output phase.

- Collect the message $(\mathsf{msg}_i^\ell, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma_i^\ell, \mathsf{vk}_i, \mathsf{cert}_i)$ for round ℓ from each party $p_i \in [n] \backslash \{j\}$ and send $(\mathsf{NextMsg}, \mathsf{sid}, \{\mathsf{msg}_i^\ell, \sigma_i^\ell, \mathsf{vk}_i, \mathsf{cert}_i\}_{i \in [n]})$ to $T_j^{\text{HT-FBS}}$.
- Upon receiving y_j from $T_j^{\text{HT-FBS}}$ output it.

Check-channel. The party p_j inspects all messages that are sent on the channel. If a message $(m, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma, \mathsf{vk}, \mathsf{cert})$ is received from a party p_i check if $\mathsf{Ver}(\mathsf{vk}_i, m || \Pi^{\mathsf{MPC}} || \ell, \sigma) = 1$ and $\mathsf{Ver}(\mathsf{mpk}, \mathsf{vk} || \mathsf{sid} || i, \mathsf{cert}) = 1$ for some $\ell \in [m]$. If it is not, then output $(\bot, p_i)$ and stop.

Fig. 5. Protocol $\Pi^{\text{HT-FBS}}$ followed by p_j to achieve fallback security.

from the tokens are encrypted and authenticated. If tokens are in addition compromised, then our fallback protocol (when based on a MPC protocol with (G)UC security in presence of a malicious broadcast), also preserves security.

VI. CP MPC FOR RATIONAL ADVERSARIES

We start this section by using the RPD framework to study the feasibility of implementing functionalities $\mathcal{F}$ in a collusion preserving way, against incentive-driven attackers that may even choose to abort the protocol. We extend the RPD model to the case of collusion-preserving functionalities (which we call *RPD-CP*). We prove that our protocols Π^{HT} and $\Pi^{\text{HT-FBS}}$ disincentivize collusion in this model, when there is a sufficient penalty to the attacker for aborting. Recall that Π^{HT} and $\Pi^{\text{HT-FBS}}$ are CP against non-aborting adversaries, and can (publicly) identify a corrupt party in case of an abort. Finally, we show how this model can be applied in practice, e.g. using the blockchain. In particular, we can abstract the blockchain by means of an ideal functionality that allows the parties to deposit collateral, which can be reclaimed on the agreement of all parties. We provide a protocol Π_{pen} that uses this functionality to concretely penalize an attacker for aborting Π^{HT} or $\Pi^{\text{HT-FBS}}$.

A. RPD-CP: Tuning RPD to the CP setting

a) The Attack Model: Following the RPD framework [28], we capture collusion-preservation against incentive-driven attackers, by considering attacks as part of an *attack game* $\mathcal{G}_{\mathcal{M}}$ between a protocol designer $\mathtt{D}$ and attacker $\mathtt{A}$. Here, $\mathtt{D}$ comes up with a protocol Π, and the attacker $\mathtt{A} \in \mathtt{ITM}$ generates a set of adversaries/adversarial strategies

The functionality interacts with a set of parties $\mathcal{P} = \{p_1, \ldots, p_n\}$. It maintains a set of honest parties $\mathcal{H} \subseteq \mathcal{P}$, and a set of malicious parties $\mathcal{I} \subseteq \mathcal{P}$

- Upon receiving $(\mathtt{COLLUDE}, \mathsf{sid}, m)$ from party $p_i \in \mathcal{I}$, send message $(\mathtt{SUB\_MSG}, \mathsf{sid}, p_i, m)$ to p_j, for all $p_j \in \mathcal{P} - \{p_i\}$.
- Upon receiving $(\mathtt{ABORT}, \mathsf{sid})$ from a party p_i, send $(\mathtt{ABORT}, p_i)$ to p_j, for all $p_j \in \mathcal{P}$-$\{p_i\}$ and stop.
- Upon receiving a message that is consistent with the interface of $\mathcal{F}$ act as $\mathcal{F}$ would do acting as a proxy between $\mathcal{F}$ and the parties in $\mathcal{P}$.

Fig. 6. $\langle \mathcal{F} \rangle$ weakens $\mathcal{F}$ with commands $\mathtt{COLLUDE}$ and $\mathtt{ABORT}$.

$\mathtt{A}(\Pi) = \mathcal{A} = \{\mathcal{A}_i\}_{i \in \mathcal{I}}$, $\mathcal{I} \subseteq [n]$ to attack it. The attacker's utility $u_{\mathtt{A}}$ is then a function of the choice of protocol Π and adversarial strategies $\mathcal{A}$. The attack model $\mathcal{M} = (\mathcal{F}, \langle \mathcal{F} \rangle, v_{\mathtt{A}})$ (and $v_{\mathtt{D}}$, if we also consider the designer's utility) encompasses the parameters in this game—$\langle \mathcal{F} \rangle$ is the weaker version of the functionality $\mathcal{F}$ we wish to implement. $\langle \mathcal{F} \rangle$ explicitly allows 1) CP to be broken by sending a colluding message to other adversarial parties and 2) the adversarial parties to abort and being identified by all the other parties that are running the protocol. Note that in contrast to monolithic adversaries and simulators, in CP the ideal adversarial parties do not automatically share their views and must use $\langle \mathcal{F} \rangle$ to collude (see. Fig. 6). Lastly, the attacker's utility $u_{\mathtt{A}}$ is defined based on a value function $v_{\mathtt{A}}$, which assigns payoffs to events occurring in the ideal world—more details below.

1) Utility of the attacker $\mathtt{A}$*:* The utility of the attacker $u_{\mathtt{A}}$ is a function mapping protocols and sets of adversaries, i.e. the strategy profile $(\Pi, \mathcal{A})$, to a real number. In our case, utility depends on whether a set of simulators must collude via a weakness in $\langle \mathcal{F} \rangle$ in order to emulate $\mathcal{A}$ in Π, and whether the simulators trigger an abort. More formally: First, we have a value function $v_{\mathtt{A}}$, defined in the attack model, which maps the views of the simulators and environment in the ideal world to a real value. Then, we define the real payoff of a particular $\mathcal{A}$ attacking the protocol, as the minimum payoff over all simulators that can emulate $\mathcal{A}$. Finally, $u_{\mathtt{A}}(\Pi, \mathcal{A})$ is the real payoff of $\mathcal{A}$, maximized over all possible environments $\mathcal{Z}$.

a) Ideal payoff of a set of simulators: In more detail, we define the real-valued random variable ensemble $\{v_{\mathtt{A}}^{\langle \mathcal{F} \rangle, \mathcal{S}, \mathcal{Z}}(k, z)\}_{k \in \mathbb{N}, z \in \{0,1\}^*}$ (or $v_{\mathtt{A}}^{\langle \mathcal{F} \rangle, \mathcal{S}, \mathcal{Z}}$ for short) as the random variable ensemble resulting from applying value function $v_{\mathtt{A}}$ to the view of the environment $\mathcal{Z}$ and a set of simulators $\mathcal{S} = \{\mathcal{S}_i\}_{i \in \mathcal{I}}$ in the ideal execution. The ideal expected payoff of a particular set of simulators $\mathcal{S}$ with respect to $\mathcal{Z}$ is defined as the expected value: $U_{I^{\mathtt{A}}}^{\langle \mathcal{F} \rangle}(\mathcal{S}, \mathcal{Z}) = E(v_{\mathtt{A}}^{\langle \mathcal{F} \rangle, \mathcal{S}, \mathcal{Z}})$.

b) Real payoff of a set of adversaries: Recall that given a setup $\bar{\mathcal{G}}$ and resource $\mathcal{R}$, a $\{\bar{\mathcal{G}}, \mathcal{R}\}$-exclusive (that is, the protocol only uses $\bar{\mathcal{G}}, \mathcal{R}$) protocol Π realizes a CP-functionality $\langle \mathcal{F} \rangle$ if, for all $\mathcal{I} \subseteq [n]$, and independent (rather than monolithic) adversaries $\mathcal{A} = \{\mathcal{A}_i\}_{i \in \mathcal{I}}$, there exists a collection of efficiently computable transformations from ITMs to ITMs $\mathsf{Sim} = \{\mathsf{Sim}_i\}_{i \in \mathcal{I}}$ such that the simulator $\mathcal{S}_i = \mathsf{Sim}_i(\mathcal{A}_i)$

emulates $\mathcal{A}_i$. That is, the environment $\mathcal{Z}$ cannot distinguish between the real world with $\mathcal{A}$ and resource $\mathcal{R}$, and ideal world with $\mathcal{S} = \{\mathsf{Sim}_i(\mathcal{A}_i)\}_{i\in\mathcal{I}}$ and $\langle\mathcal{F}\rangle$. Let $\langle\mathcal{F}\rangle$ be a CP functionality and Π be a protocol. Denote $\mathcal{C}_\mathcal{A}$ as the class of simulators $\mathcal{S} = \{\mathcal{S}_i\}_{i\in\mathcal{I}}$ that can emulate the adversarial parties $\mathcal{A} = \{\mathcal{A}_i\}_{i\in\mathcal{I}}$ for $\mathcal{I} \subseteq [n]$. That is, for setup $\bar{\mathcal{G}}$ and resource $\mathcal{R}$, $\mathcal{C}_\mathcal{A} = \Big\{\mathsf{Sim}(\mathcal{A}) = \{\mathsf{Sim}_i(\mathcal{A}_i)\}_{i\in\mathcal{I}} \mid \forall i \in \mathcal{I}$: Sim_i an efficiently computable mapping from ITM to ITM, $\forall\mathcal{Z}$: CP-EXEC$^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi,\mathcal{A},\mathcal{Z}} \approx$ CP-EXEC$^{\bar{\mathcal{G}},\langle\mathcal{F}\rangle}_{\Pi,\mathsf{Sim}(\mathcal{A}),\mathcal{Z}}\Big\}$. The expected payoff of a set of adversaries and environment $(\mathcal{A},\mathcal{Z})$ is then defined as $U_\mathtt{A}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A},\mathcal{Z}) = \inf_{\mathcal{S}\in\mathcal{C}_\mathcal{A}}\{U_{I^\mathtt{A}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z})\}$. The attacker's utility is then maximized over all environments $\mathcal{Z}$, i.e., $u_\mathtt{A}(\Pi,\mathcal{A}) := \hat{U}_\mathtt{A}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A}) = \sup_{\mathcal{Z}\in\mathtt{ITM}}\{U_\mathtt{A}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A},\mathcal{Z})\}$.

2) Utility of the protocol designer D*:* In [28], the attack game is assumed to be zero-sum, i.e. the designer's utility $u_\mathtt{D} = -u_\mathtt{A}$. To remove this assumption, we follow the methodology of a more recent work [6] to define $u_\mathtt{D}$. In more detail, for each $(\Pi,\mathcal{A})$, we must assign utility for the designer using the *same* simulators and environments as those used for the attacker. Let $\mathcal{S}_\mathtt{A}$ denote the class of simulators that were used to obtain the utility of the attacker, and $\mathcal{Z}_\mathtt{A}$ denote the class of environments maximizing the utility for simulators in $\mathcal{S}_\mathtt{A}$. That is, $\mathcal{S}_\mathtt{A} = \Big\{\mathcal{S}\in\mathcal{C}_\mathcal{A}\colon \sup_{\mathcal{Z}\in\mathtt{ITM}}\{U_{I^\mathtt{A}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z})\} = u_\mathtt{A}(\Pi,\mathcal{A})\Big\}$ and $\mathcal{Z}_\mathtt{A} = \Big\{\mathcal{Z}\in\mathtt{ITM}\colon \text{for some } \mathcal{S}\in\mathcal{S}_\mathtt{A}\ ,\ U_{I^\mathtt{A}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z}) = u_\mathtt{A}(\Pi,\mathcal{A})\Big\}$.

Then, let $v_\mathtt{D}^{\langle\mathcal{F}\rangle,\mathcal{S},\mathcal{Z}}$ and $U_{I^\mathtt{D}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z})$ be defined similar to the payoffs $v_\mathtt{A}^{\langle\mathcal{F}\rangle,\mathcal{S},\mathcal{Z}}$ and $U_{I^\mathtt{A}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z})$ respectively. Again following the definitions of [6], the real payoff of the designer is $U_\mathtt{D}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A},\mathcal{Z}) = \sup_{\mathcal{S}\in\mathcal{S}_\mathtt{A}}\{U_{I^\mathtt{D}}^{\langle\mathcal{F}\rangle}(\mathcal{S},\mathcal{Z})\}$. The utility of the designer is then $u_\mathtt{D}(\Pi,\mathcal{A}) := \hat{U}_\mathtt{D}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A}) = \inf_{\mathcal{Z}\in\mathcal{Z}_\mathtt{A}}\{U_\mathtt{D}^{\Pi,\langle\mathcal{F}\rangle}(\mathcal{A},\mathcal{Z})\}$ We can extend the attack model with the value function of the designer $v_\mathtt{D}$: $\mathcal{M} = (\mathcal{F},\langle\mathcal{F}\rangle, v_\mathtt{A}, v_\mathtt{D})$.

3) Attack-payoff security with collusion-preservation: Similar to the definition of *attack-payoff secure* in [6], [28], we define collusion preserving attack payoff (CPAP). Intuitively, a protocol is CPAP with respect to an attack model $\mathcal{M} = (\mathcal{F},\langle\mathcal{F}\rangle, v_\mathtt{A})$ if it enjoys security and collusion-preservation under this model. That is, no attacker can gain more utility from running our protocol, than running the dummy protocol that uses a functionality $\mathcal{F}$ as a resource.

Definition 2 (CPAP)**.** *Let $\mathcal{M} = (\mathcal{F},\langle\mathcal{F}\rangle, v_\mathtt{A})$ be an attack model and Π a $\{\bar{\mathcal{G}},\mathcal{R}\}$-exclusive protocol that realizes $\langle\mathcal{F}\rangle$. We say that Π is* CPAP *in $\mathcal{M}$ if $\sup_{\mathcal{A}\in\mathtt{ITM}} u_\mathtt{A}(\Pi,\mathcal{A}) \overset{negl}{\leq} \sup_{\mathcal{A}\in\mathtt{ITM}} u_\mathtt{A}(\Phi^\mathcal{F},\mathcal{A})$ where $\Phi^\mathcal{F}$ is the dummy $\{\bar{\mathcal{G}},\mathcal{F}\}$-hybrid protocol which forwards all inputs to and outputs from functionality $\mathcal{F}$.*

To complete our framework, we also define ϵ-subgame-perfect equilibrium from [28], and define CPIC similarly to the definition of *incentive compatible (IC)* in [6]. Informally, a strategy profile is an ϵ-subgame-perfect equilibrium if no deviation could improve utilities by more than ϵ. Intuitively, a protocol Π is incentive compatible when even the designer is incentivized to stick with it.

Definition 3 (ϵ-subgame-perfect equilibrium [28])**.** *Let $\mathcal{G}_\mathcal{M}$ be an attack game. A strategy profile $(\Pi,\mathtt{A}(\Pi))$ is an* ϵ-subgame perfect equilibrium *in $\mathcal{G}_\mathcal{M}$ if the following conditions hold: (1) for any $\Pi'\in\mathtt{ITM}^n$, $u_\mathtt{D}(\Pi',\mathtt{A}(\Pi')) \leq u_\mathtt{D}(\Pi,\mathtt{A}(\Pi)) + \epsilon$, and (2) for any $\mathtt{A}'\in\mathtt{ITM}$, $u_\mathtt{A}(\Pi,\mathtt{A}'(\Pi)) \leq u_\mathtt{A}(\Pi,\mathtt{A}(\Pi)) + \epsilon$.*

Definition 4 (CPIC)**.** *Let Π be a $\{\bar{\mathcal{G}},\mathcal{R}\}$-exclusive protocol and $\boldsymbol{\Pi}$ be a set of polynomial-time $\{\bar{\mathcal{G}},\mathcal{R}\}$-exclusive protocols. We say that Π is* $\boldsymbol{\Pi}$-CPIC *in the attack model $\mathcal{M}$ iff for some $\mathtt{A}\in\mathtt{ITM}$, $(\Pi,\mathtt{A}(\Pi))$ is a $\nu(\lambda)$-subgame perfect equilibrium on the restricted attack game where the set of deviations of the designer is $\boldsymbol{\Pi}$.*

B. $\Pi^{\mathtt{HT}}$ and $\Pi^{\mathtt{HT\text{-}FBS}}$ with incentives

In this section we show that the protocols $\Pi^{\mathtt{HT}}$ and $\Pi^{\mathtt{HT\text{-}FBS}}$ presented in the previous sections are *CPAP* for a natural class of utilities $u_\mathtt{A}$. For simplicity we consider the ideal functionality $\mathcal{F}^f$ for SFE parameterized by the function $f:\mathbb{F}^n\to\mathbb{F}$ (this form is without loss of generality—see, e.g., [46]), which we assume is a CP-well-formed functionality. We refer the reader to App. H for a formal definition of $\mathcal{F}^f$. We consider the following events the value function $v_\mathtt{A}$ is concerned with. These are events defined on the views of the environment, the (relaxed) CP-functionality $\langle\mathcal{F}^f\rangle$, and the simulators $\mathcal{S} = \{\mathcal{S}_i\}_{i\in\mathcal{I}}$, given adversaries $\mathcal{A} = \{\mathcal{A}_i\}_{i\in\mathcal{I}}$:

- Define the event $E_\mathtt{collude}$ as follows: For some $i\in\mathcal{I}$ and message m, the ith simulator $\mathcal{S}_i$ sends the message $(\mathtt{COLLUDE},\mathsf{sid},m)$ to $\langle\mathcal{F}^f\rangle$.
- Define the event $E_\mathtt{abort}$ as follows: For some $i\in\mathcal{I}$, party p_i aborts and is identified by all the parties as having aborted.

Now, we define the payoffs assigned by $v_\mathtt{A}$ to the events above. Denote by $\gamma_\mathtt{collude}$ the utility for the attacker obtained by triggering $E_\mathtt{collude}$. Denote by $\gamma_\mathtt{abort}$ the penalty incurred as result of a malicious party being identified by the honest parties as an aborting party. Then, the utility of the attacker is: $u_\mathtt{A}(\Pi,\mathcal{A}) = \sup_\mathcal{Z}\left\{\inf_{\mathcal{S}\in\mathcal{C}_\mathcal{A}}\left\{\gamma_\mathtt{collude}\Pr[E_\mathtt{collude}] - \gamma_\mathtt{abort}\Pr[E_\mathtt{abort}]\right\}\right\}$

Our protocol satisfies CPAP security under the condition that the penalty of being identified as having aborted is greater than the gain from sending a colluding message. In Sec. VI-C, we will discuss a penalization scheme that makes these penalties concrete.

Theorem 3. *Let $\mathcal{F}^f$ be an ideal CP-well-formed functionality, and $\langle\mathcal{F}^f\rangle$ be as defined in Fig. 6. Let $v_\mathtt{A}$ be as defined above, for any $\gamma_\mathtt{collude}$ and $\gamma_\mathtt{abort}$ such that $\gamma_\mathtt{abort} > \gamma_\mathtt{collude}$. Then the protocol $\Pi^{\mathtt{HT}}$ described in Sec. V (and the protocol $\Pi^{\mathtt{HT\text{-}FBS}}$ described in Sec. V-A) is* CPAP secure *in the attack model $\mathcal{M} = (\mathcal{F}^f,\langle\mathcal{F}^f\rangle, v_\mathtt{A})$.*

Proof sketch (full proof see App. B). To prove this theorem we rely on the observation that $\Pi^{\mathtt{HT}}$ ($\Pi^{\mathtt{HT\text{-}FBS}}$) is collusion-preserving for $\mathcal{F}^f$ as long as nobody aborts. Moreover, $\Pi^{\mathtt{HT}}$ and $\Pi^{\mathtt{HT\text{-}FBS}}$ achieve (publicly) identifiable abort, since the only way subliminally communicate is by sending a message which is incompatible with the protocol description. Given the way we have set the payoffs, it is always inconvenient for the

adversary to trigger the collusion event $E_{\texttt{collude}}$, as it causes the abort event $E_{\texttt{abort}}$.

C. Realizing the incentives

The goal of this section is to create a penalization scheme that translates the utilities defined above to concrete (monetary) values. Below, we describe a simple protocol Π_{pen} that realizes a CP functionality $\mathcal{F}$ assuming no aborts, and disincentivizes aborts with penalization. As we show in Theorem 4 below, Π_{pen} achieves CPAP, and is incentive compatible for the designer (i.e. is CPIC), assuming the existence of one honest party. The protocol assumes a functionality $\mathcal{F}_{\text{pen}}$, which allows each party to deposit a collateral of amount d (a parameter of the functionality). Parties can reclaim their collateral *if and only if* all parties send the functionality a RECLAIM message (for more detail on $\mathcal{F}_{\text{pen}}$, see App. F). In Π_{pen}, honest parties only send RECLAIM if they do not detect an abort during execution of $\Pi^{\texttt{HT}}$ or $\Pi^{\texttt{HT-FBS}}$. The protocol Π_{pen} works as follows.

1) Each party sends (DEPOSIT) to $\mathcal{F}_{\text{pen}}$, to deposit a collateral of amount d. If the functionality returns (DEPOSIT, 1), proceed. Otherwise, stop.
2) Run $\Pi^{\texttt{HT}}$ (resp. $\Pi^{\texttt{HT-FBS}}$) (possibly multiple times for reactive functionality, using secret sharing to maintain secret intermediate state) on the broadcast channel $\mathcal{B}$.
3) If $\Pi^{\texttt{HT}}$ (resp. $\Pi^{\texttt{HT-FBS}}$) did not abort, then all parties send (RECLAIM) to $\mathcal{F}_{\text{pen}}$. $\mathcal{F}_{\text{pen}}$ ensures that parties receive their deposits back if and only if everyone sends (RECLAIM).

We define below a class of utilities for the above protocol, making the arguably natural assumption that the attacker (who chooses the adversarial strategies for all corrupted parties) cares about the sum of deposits lost/gained by all corrupt parties, and similarly, the protocol designer cares about the deposits of all honest parties. Let $E_{\texttt{collude}}$ be the event the simulator sends a COLLUDE message in $\langle\mathcal{F}\rangle$; $E^{\texttt{A}}_{\texttt{deposit}}(t)$ (resp. $E^{\texttt{D}}_{\texttt{deposit}}(m)$) is the event t corrupt (resp. m honest) parties send (DEPOSIT) to $\mathcal{F}_{\text{pen}}$ and $\mathcal{F}_{\text{pen}}$ returns (DEPOSIT, 1); $E^{\texttt{A}}_{\texttt{reclaim}}(t)$ (resp. $E^{\texttt{D}}_{\texttt{reclaim}}(m)$) is the event where t corrupt (resp. m honest) parties receive the message (RECLAIM, 1). Sets $\mathcal{C}_{\mathcal{A}}, \mathcal{S}_{\texttt{A}}, \mathcal{Z}_{\texttt{A}}$ are defined in Section VI-A.

$$u_{\texttt{A}}(\Pi,\mathcal{A}) = \sup_{\mathcal{Z}}\Big\{\inf_{\mathcal{S}\in\mathcal{C}_{\mathcal{A}}}\Big\{\gamma_{\texttt{collude}}\Pr[E_{\texttt{collude}}] - \sum_{t\in[n]} td\cdot\Pr[E^{\texttt{A}}_{\texttt{deposit}}(t)] + \sum_{t\in[n]} td\cdot\Pr[E^{\texttt{A}}_{\texttt{reclaim}}(t)]\Big\}\Big\}$$

$$u_{\texttt{D}}(\Pi,\mathcal{A}) = \inf_{\mathcal{Z}\in\mathcal{Z}_{\texttt{A}}}\Big\{\sup_{\mathcal{S}\in\mathcal{S}_{\texttt{A}}}\Big\{-\gamma_{\texttt{collude}}\Pr[E_{\texttt{collude}}] - \sum_{m\in[n]} md\cdot\Pr[E^{\texttt{D}}_{\texttt{deposit}}(m)] + \sum_{m\in[n]} md\cdot\Pr[E^{\texttt{D}}_{\texttt{reclaim}}(m)]\Big\}\Big\}$$

Informally, we show CPIC and CPAP by proving that the strategy profile $(\Pi_{\text{pen}}, \mathcal{A})$, where the $\mathcal{A}$ is the passive adversarial strategy that just follows Π_{pen} (and does not collude/abort), is an equilibrium solution. We observe that corrupt parties know that if they try to gain more utility by colluding (and thus abort), there is at least one honest party to ensure they lose their collateral, which means they have no incentive to deviate from Π_{pen}. The protocol designer also has no incentive to deviate from Π_{pen} if the adversary is passive, which means this strategy profile is an equilibrium.

Theorem 4 (Proof in App. B)**.** *Let $\mathcal{F}$ be an ideal CP-well-formed functionality, and $\langle\mathcal{F}\rangle$ be as defined in Fig. 6. Let $v_{\texttt{A}}$ and $v_{\texttt{D}}$ be defined in the utility functions above, let $td > \gamma_{\texttt{collude}}$ where t is the number of corrupt parties, and assume there is at least one honest party. Then, the $\{\bar{\mathcal{G}},\mathcal{R}\}$-exclusive protocol Π_{pen}, where $\mathcal{R}$ is broadcast and $\bar{\mathcal{G}} = T^{\texttt{HT}}, \mathcal{F}_{pen}$ (resp. $\bar{\mathcal{G}} = T^{\texttt{HT-FBS}}, \mathcal{F}_{pen}$) :*

- *collusion-preservingly emulates the $\{\bar{\mathcal{G}},\mathcal{F}\}$-exclusive protocol ϕ assuming the adversary does not abort,*
- *is CPAP secure in the attack model $\mathcal{M}$, and*
- *is $\mathbf{\Pi}$-CPIC in $\mathcal{M}$.*

Where $\mathcal{M} = (\mathcal{F}, \langle\mathcal{F}\rangle, v_{\texttt{A}}, v_{\texttt{D}})$ and $\mathbf{\Pi}$ is the set of all protocols.

One could implement deposit/reclaim via, e.g., a smart contract-enabled blockchain. We refer to App. D-A for further discussion on penalization and on how to avoid that also honest parties are penalized in the case a misbehavior is detected.

REFERENCES

[1] J. Alwen, J. Katz, Y. Lindell, G. Persiano, a. shelat, and I. Visconti. Collusion-free multiparty computation in the mediated model. In S. Halevi, editor, *Advances in Cryptology – CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 524–540, Santa Barbara, CA, USA, Aug. 16–20, 2009. Springer, Heidelberg, Germany.

[2] J. Alwen, J. Katz, U. Maurer, and V. Zikas. Collusion-preserving computation. In R. Safavi-Naini and R. Canetti, editors, *Advances in Cryptology – CRYPTO 2012*, volume 7417 of *Lecture Notes in Computer Science*, pages 124–143, Santa Barbara, CA, USA, Aug. 19–23, 2012. Springer, Heidelberg, Germany.

[3] J. Alwen, R. Ostrovsky, H.-S. Zhou, and V. Zikas. Incoercible multiparty computation and universally composable receipt-free voting. In R. Gennaro and M. J. B. Robshaw, editors, *Advances in Cryptology – CRYPTO 2015, Part II*, volume 9216 of *Lecture Notes in Computer Science*, pages 763–780, Santa Barbara, CA, USA, Aug. 16–20, 2015. Springer, Heidelberg, Germany.

[4] J. Alwen, a. shelat, and I. Visconti. Collusion-free protocols in the mediated model. In D. Wagner, editor, *Advances in Cryptology – CRYPTO 2008*, volume 5157 of *Lecture Notes in Computer Science*, pages 497–514, Santa Barbara, CA, USA, Aug. 17–21, 2008. Springer, Heidelberg, Germany.

[5] M. Andrychowicz, S. Dziembowski, D. Malinowski, and L. Mazurek. Secure multiparty computations on bitcoin. In *2014 IEEE Symposium on Security and Privacy*, pages 443–458, Berkeley, CA, USA, May 18–21, 2014. IEEE Computer Society Press.

[6] C. Badertscher, J. A. Garay, U. Maurer, D. Tschudi, and V. Zikas. But why does it work? A rational protocol design treatment of bitcoin. In J. B. Nielsen and V. Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2018, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 34–65, Tel Aviv, Israel, Apr. 29 – May 3, 2018. Springer, Heidelberg, Germany.

[7] C. Badertscher, U. Maurer, D. Tschudi, and V. Zikas. Bitcoin as a transaction ledger: A composable treatment. In J. Katz and H. Shacham, editors, *Advances in Cryptology – CRYPTO 2017, Part I*, volume 10401 of *Lecture Notes in Computer Science*, pages 324–356, Santa Barbara, CA, USA, Aug. 20–24, 2017. Springer, Heidelberg, Germany.

[8] R. Bahmani, M. Barbosa, F. Brasser, B. Portela, A.-R. Sadeghi, G. Scerri, and B. Warinschi. Secure multiparty computation from SGX. In A. Kiayias, editor, *FC 2017: 21st International Conference on Financial Cryptography and Data Security*, volume 10322 of *Lecture Notes in Computer Science*, pages 477–497, Sliema, Malta, Apr. 3–7, 2017. Springer, Heidelberg, Germany.

[9] M. Barbosa, B. Portela, G. Scerri, and B. Warinschi. Foundations of hardware-based attested computation and application to sgx. In *2016 IEEE European Symposium on Security and Privacy (EuroSP)*, pages 245–260, Los Alamitos, CA, USA, mar 2016. IEEE Computer Society.

[10] F. Benhamouda and H. Lin. k-round multiparty computation from k-round oblivious transfer via garbled interactive circuits. In J. B. Nielsen and V. Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2018, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 500–532, Tel Aviv, Israel, Apr. 29 – May 3, 2018. Springer, Heidelberg, Germany.

[11] I. Bentov and R. Kumaresan. How to use bitcoin to design fair protocols. In J. A. Garay and R. Gennaro, editors, *Advances in Cryptology – CRYPTO 2014, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 421–439, Santa Barbara, CA, USA, Aug. 17–21, 2014. Springer, Heidelberg, Germany.

[12] I. Bentov, R. Kumaresan, and A. Miller. Instantaneous decentralized poker. In T. Takagi and T. Peyrin, editors, *Advances in Cryptology – ASIACRYPT 2017, Part II*, volume 10625 of *Lecture Notes in Computer Science*, pages 410–440, Hong Kong, China, Dec. 3–7, 2017. Springer, Heidelberg, Germany.

[13] B. H. Bloom. Space/time trade-offs in hash coding with allowable errors. *Commun. ACM*, 13(7):422426, July 1970.

[14] R. Canetti. Security and composition of multiparty cryptographic protocols. *Journal of Cryptology*, 13(1):143–202, Jan. 2000.

[15] R. Canetti. Universally composable security: A new paradigm for cryptographic protocols. Cryptology ePrint Archive, Report 2000/067, 2000. http://eprint.iacr.org/2000/067.

[16] R. Canetti. Universally composable signatures, certification and authentication. Cryptology ePrint Archive, Report 2003/239, 2003. http://eprint.iacr.org/2003/239.

[17] R. Canetti, Y. Dodis, R. Pass, and S. Walfish. Universally composable security with global setup. In S. P. Vadhan, editor, *TCC 2007: 4th Theory of Cryptography Conference*, volume 4392 of *Lecture Notes in Computer Science*, pages 61–85, Amsterdam, The Netherlands, Feb. 21–24, 2007. Springer, Heidelberg, Germany.

[18] R. Canetti, A. Jain, and A. Scafuro. Practical UC security with a global random oracle. In G.-J. Ahn, M. Yung, and N. Li, editors, *ACM CCS 2014: 21st Conference on Computer and Communications Security*, pages 597–608, Scottsdale, AZ, USA, Nov. 3–7, 2014. ACM Press.

[19] R. Canetti, Y. Lindell, R. Ostrovsky, and A. Sahai. Universally composable two-party and multi-party secure computation. In *34th Annual ACM Symposium on Theory of Computing*, pages 494–503, Montréal, Québec, Canada, May 19–21, 2002. ACM Press.

[20] R. Canetti and M. Vald. Universally composable security with local adversaries. In I. Visconti and R. D. Prisco, editors, *SCN 12: 8th International Conference on Security in Communication Networks*, volume 7485 of *Lecture Notes in Computer Science*, pages 281–301, Amalfi, Italy, Sept. 5–7, 2012. Springer, Heidelberg, Germany.

[21] R. Cohen, J. A. Garay, and V. Zikas. Broadcast-optimal two-round MPC. In A. Canteaut and Y. Ishai, editors, *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part II*, volume 12106 of *Lecture Notes in Computer Science*, pages 828–858. Springer, 2020.

[22] V. Costan and S. Devadas. Intel SGX explained. Cryptology ePrint Archive, Report 2016/086, 2016. http://eprint.iacr.org/2016/086.

[23] I. Damgård, J. B. Nielsen, and D. Wichs. Universally composable multiparty computation with partially isolated parties. In O. Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 315–331. Springer, Heidelberg, Germany, Mar. 15–17, 2009.

[24] D. Dolev and H. R. Strong. Authenticated algorithms for byzantine agreement. *SIAM Journal on Computing*, 12(4):656–666, 1983.

[25] N. Döttling, D. Kraschewski, and J. Müller-Quade. Unconditional and composable security using a single stateful tamper-proof hardware token. In Y. Ishai, editor, *TCC 2011: 8th Theory of Cryptography Conference*, volume 6597 of *Lecture Notes in Computer Science*, pages 164–181, Providence, RI, USA, Mar. 28–30, 2011. Springer, Heidelberg, Germany.

[26] B. Fisch, D. Vinayagamurthy, D. Boneh, and S. Gorbunov. IRON: Functional encryption using intel SGX. In B. M. Thuraisingham, D. Evans, T. Malkin, and D. Xu, editors, *ACM CCS 2017: 24th Conference on Computer and Communications Security*, pages 765–782, Dallas, TX, USA, Oct. 31 – Nov. 2, 2017. ACM Press.

[27] J. A. Garay, J. Katz, R. Kumaresan, and H.-S. Zhou. Adaptively secure broadcast, revisited. In C. Gavoille and P. Fraigniaud, editors, *30th ACM Symposium Annual on Principles of Distributed Computing*, pages 179–186, San Jose, CA, USA, June 6–8, 2011. Association for Computing Machinery.

[28] J. A. Garay, J. Katz, U. Maurer, B. Tackmann, and V. Zikas. Rational protocol design: Cryptography against incentive-driven adversaries. In *54th Annual Symposium on Foundations of Computer Science*, pages 648–657, Berkeley, CA, USA, Oct. 26–29, 2013. IEEE Computer Society Press.

[29] O. Goldreich. *Foundations of cryptography: volume 2, basic applications*. Cambridge university press, 2009.

[30] S. Goldwasser and M. Bellare. Lecture notes on cryptography. *Summer course Cryptography and computer security at MIT*, 1999:1999, 1996.

[31] S. Goldwasser, Y. T. Kalai, and G. N. Rothblum. One-time programs. In D. Wagner, editor, *Advances in Cryptology – CRYPTO 2008*, volume 5157 of *Lecture Notes in Computer Science*, pages 39–56, Santa Barbara, CA, USA, Aug. 17–21, 2008. Springer, Heidelberg, Germany.

[32] S. Goldwasser and S. Micali. Probabilistic encryption and how to play mental poker keeping secret all partial information. In *14th Annual ACM Symposium on Theory of Computing*, pages 365–377, San Francisco, CA, USA, May 5–7, 1982. ACM Press.

[33] S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.

[34] V. Goyal, Y. Ishai, A. Sahai, R. Venkatesan, and A. Wadia. Founding cryptography on tamper-proof hardware tokens. In D. Micciancio, editor, *TCC 2010: 7th Theory of Cryptography Conference*, volume 5978 of *Lecture Notes in Computer Science*, pages 308–326, Zurich, Switzerland, Feb. 9–11, 2010. Springer, Heidelberg, Germany.

[35] M. Hirt and V. Zikas. Adaptively secure broadcast. In H. Gilbert, editor, *Advances in Cryptology – EUROCRYPT 2010*, volume 6110 of *Lecture Notes in Computer Science*, pages 466–485, French Riviera, May 30 – June 3, 2010. Springer, Heidelberg, Germany.

[36] Y. Ishai, R. Ostrovsky, and V. Zikas. Secure multi-party computation with identifiable abort. In J. A. Garay and R. Gennaro, editors, *Advances in Cryptology – CRYPTO 2014, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 369–386, Santa Barbara, CA, USA, Aug. 17–21, 2014. Springer, Heidelberg, Germany.

[37] J. Katz. Universally composable multi-party computation using tamper-proof hardware. In M. Naor, editor, *Advances in Cryptology – EUROCRYPT 2007*, volume 4515 of *Lecture Notes in Computer Science*, pages 115–128, Barcelona, Spain, May 20–24, 2007. Springer, Heidelberg, Germany.

[38] J. Katz, U. Maurer, B. Tackmann, and V. Zikas. Universally composable synchronous computation. In A. Sahai, editor, *TCC 2013: 10th Theory of Cryptography Conference*, volume 7785 of *Lecture Notes in Computer Science*, pages 477–498, Tokyo, Japan, Mar. 3–6, 2013. Springer, Heidelberg, Germany.

[39] A. Kiayias, H.-S. Zhou, and V. Zikas. Fair and robust multi-party computation using a global transaction ledger. In M. Fischlin and J.-S. Coron, editors, *Advances in Cryptology – EUROCRYPT 2016, Part II*, volume 9666 of *Lecture Notes in Computer Science*, pages 705–734, Vienna, Austria, May 8–12, 2016. Springer, Heidelberg, Germany.

[40] R. Kumaresan and I. Bentov. How to use bitcoin to incentivize correct computations. In G.-J. Ahn, M. Yung, and N. Li, editors, *ACM CCS 2014: 21st Conference on Computer and Communications Security*, pages 30–41, Scottsdale, AZ, USA, Nov. 3–7, 2014. ACM Press.

[41] R. Kumaresan and I. Bentov. Amortizing secure computation with penalties. In E. R. Weippl, S. Katzenbeisser, C. Kruegel, A. C. Myers, and S. Halevi, editors, *ACM CCS 2016: 23rd Conference on Computer and Communications Security*, pages 418–429, Vienna, Austria, Oct. 24–28, 2016. ACM Press.

[42] R. Kumaresan, T. Moran, and I. Bentov. How to use bitcoin to play decentralized poker. In I. Ray, N. Li, and C. Kruegel, editors, *ACM CCS 2015: 22nd Conference on Computer and Communications Security*, pages 195–206, Denver, CO, USA, Oct. 12–16, 2015. ACM Press.

[43] L. Lamport. Password authentication with insecure communication. *Commun. ACM*, 24(11):770772, Nov. 1981.

[44] M. Lepinski, S. Micali, C. Peikert, and a. shelat. Completely fair sfe and coalition-safe cheap talk. In S. Chaudhuri and S. Kutten, editors, *23rd ACM Symposium Annual on Principles of Distributed Computing*, pages 1–10, St. John's, Newfoundland, Canada, July 25–28, 2004. Association for Computing Machinery.

[45] M. Lepinski, S. Micali, and a. shelat. Collusion-free protocols. In H. N. Gabow and R. Fagin, editors, *37th Annual ACM Symposium on Theory of Computing*, pages 543–552, Baltimore, MA, USA, May 22–24, 2005. ACM Press.

[46] 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.

[47] T. Moran and G. Segev. David and Goliath commitments: UC computation for asymmetric parties using tamper-proof hardware. In N. P. Smart, editor, *Advances in Cryptology – EUROCRYPT 2008*, volume 4965 of *Lecture Notes in Computer Science*, pages 527–544, Istanbul, Turkey, Apr. 13–17, 2008. Springer, Heidelberg, Germany.

[48] J. B. Nielsen. *On protocol security in the cryptographic model*. Citeseer, 2003.

[49] T. Nilges. *The cryptographic strength of tamper-proof hardware*. PhD thesis, Karlsruhe Institute of Technology, 2015.

[50] R. Pass, E. Shi, and F. Tramèr. Formal abstractions for attested execution secure processors. In J. Coron and J. B. Nielsen, editors, *Advances in Cryptology – EUROCRYPT 2017, Part I*, volume 10210 of *Lecture Notes in Computer Science*, pages 260–289, Paris, France, Apr. 30 – May 4, 2017. Springer, Heidelberg, Germany.

[51] A. Patra and D. Ravi. On the exact round complexity of secure three-party computation. In H. Shacham and A. Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018, Part II*, volume 10992 of *Lecture Notes in Computer Science*, pages 425–458, Santa Barbara, CA, USA, Aug. 19–23, 2018. Springer, Heidelberg, Germany.

h

APPENDIX A
HARDWARE TOKENS AND BROADCAST VS. MEDIATORS

In this section, we motivate our study of collusion-preservation (CP) using authenticated broadcast channels with hardware tokens, instead of a mediator. We show (Theorem 5) that authenticated broadcast (Fig. 11) with hardware tokens is a *strictly weaker* set of assumptions than an honest mediator in a star-topology network (also called the *mediated model* [2]). Thus, our feasibility results are not immediately implied by existing results in the mediated model.

Perhaps more interestingly, although our assumptions are weaker, they are also better in terms of the fallback guarantees we can achieve. We show this by contrasting what happens when tokens are broken, versus when the mediator becomes corrupted. Whereas we can still achieve fallback security *with identifiable and unanimous abort* with corruptible hardware tokens, many such important properties become impossible when the mediator can be corrupted—even with honest majority. We refer to App. C for more details.

Theorem 5. *Hardware tokens with authenticated broadcast is a strictly weaker set of assumptions than an honest mediator in the mediated model.*

Proof. We prove the theorem by combining the two claims below.

Claim 1. Hardware tokens with authenticated broadcast does not imply an honest mediator in the mediated model.

Proof of claim: Let $\mathcal{F}$ be a CP functionality where for all $i = 1, 3, 5, \cdots$, the outputs of parties p_i and p_{i+1} are x_{i+1} and x_i respectively. That is, the pair of parties share their input with only each other. Now consider the case where all but one party are corrupted. Let p_i be the only corrupt party which is supposed to share inputs with the one honest party. Suppose in the real world p_i sends the input of the honest party over broadcast, as soon as p_i receives it as output. This means another corrupt party p_j now also knows the input of the honest party. However, in the ideal world, the simulator for p_j cannot simulate this part of the real-world execution, as $\mathcal{F}$ does not allow him to know the honest party's input. Thus, collusion-preservation is broken. On the other hand, there exists a CP protocol via an honest mediator in the mediated model for this adversary (e.g., the mediator simply ignores p_i's invalid message). Thus, an adversary breaking CP under the tokens/broadcast assumptions, does not break CP in the honest mediator/mediated model. This proves the claim.

Claim 2. An honest mediator in the mediated model implies hardware tokens with authenticated broadcast.

Proof of claim: An honest mediator (in a star-topology network) can be used as a trusted party to replace hardware tokens/authenticated broadcast. To emulate each party having his own token, the mediator runs an instance of the hardware token's code for each party. As the mediator is assumed to be honest, it is able to emulate any code correctly, while maintaining each token's secret state. To emulate authenticated broadcast, the honest mediator simply sends each message it receives to every other party with the identity of the sender attached.

□

APPENDIX B
SECURITY PROOFS

A. Proof of Theorem 1

Proof. To prove CP, we must show n independent simulators $\mathcal{S}_1, \ldots, \mathcal{S}_n$. At a very high level for all $i \in \{1, \ldots, n\}$, $\mathcal{S}_i$ can query $\mathcal{F}$ with the command $(\mathsf{GetTrapdoor}, \mathsf{sid})$. $\mathcal{F}$ checks that sid is equal to its session id, and if so, it sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to $T^{\mathtt{HT}}$. $T^{\mathtt{HT}}$ then generates the string $\widetilde{K}$, n pairs of signing-verification keys, and authenticates the verification keys using the master secret key msk. $T^{\mathtt{HT}}$ then sends this information to $\mathcal{F}$ which forwards them to the simulator $\mathcal{S}_i$. Note that if all the simulators query $\mathcal{F}$ with $(\mathsf{GetTrapdoor}, \mathsf{sid})$ they will all get the same pair of authenticated signing-verification key and the string $\widetilde{K}$. Given the above information, a simulator $\mathcal{S}_i$ can use $\widetilde{K}$ as input of a PRG to generate the randomness sufficient to: 1) to create a key $\mathsf{Kenc}^{\mathsf{sid}}$ for a secret-key encryption scheme 2) compute $n + 1$ encryptions of 0^λ and 3) authenticate the encrypted value using the appropriate signature session keys. These authenticated messages are now used to interact with $p_i^\star$ (who is not supposed to distinguish between the encryptions of 0^λ and the encryptions that contain the actual inputs of the parties). Moreover, whenever $\mathcal{S}_i$ receives the input form $p_i^\star$ he forwards it to the $\mathcal{F}$ which returns the output y_i. The crucial observation is that the messages seen by all the corrupted parties are the same since the simulators use exactly the same strategy and randomness, and since the adversary cannot forge a signature for the strong signature scheme we are using[9]. We now provide a more formal argument for the proof.

We start by assuming that at least one party is honest. In order to prove this part of the theorem we need to show a collection of efficiently computable transformations $\mathsf{Sim} = \{\mathsf{Sim}_1, \ldots, \mathsf{Sim}_n\}$ mapping ITMs to ITMs such that for every set of adversaries $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, and every PPT environment $\mathcal{Z}$ the following holds:

$$\text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi^{\mathtt{HT}},\mathcal{A},\mathcal{Z}} \approx \text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\phi,\mathsf{Sim}(\mathcal{A}),\mathcal{Z}}$$

For $i = 1, \ldots n$, the simulator $\mathcal{S}_i = \mathsf{Sim}(\mathcal{A}_i)$ queries$(\mathsf{GetTrapdoor}, \mathsf{sid})$ $\mathcal{F}^\star$ with the command $(\mathsf{GetTrapdoor}, \mathsf{sid})$. $\mathcal{F}^\star$ checks that sid is equal to its session id. If it is, then $\mathcal{F}^\star$ sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to $T^{\mathtt{HT}}$. $T^{\mathtt{HT}}$ then generates the string $\widetilde{K}$, n couples of signing-verification keys, and authenticates the verification keys using the master secret key msk (as shown in Fig 1) thus obtaining $\{\mathsf{vk}_i, \mathsf{cert}_i, \sigma_i\}_{i\in[n]}, \{e_i, \overline{y}_i\}_{i\in[n-1]}$. $\mathcal{F}^\star$ then sends $T^{\mathtt{HT}}$ this information to $\mathcal{F}^\star$ which forwards them to the simulator $\mathcal{S}_i$.

[9]We note that it is crucial to use a strong signature scheme to avoid the creation of a different valid signature for a message m from valid signatures for the same message.

From here onwards the behaviors of the simulators differ. Without loss of generality we describe how the simulator $\mathcal{S}_1$ works since $\mathcal{S}_i$ with $i \in \{2, \ldots, n-1\}$ will follow exactly the same strategy as $\mathcal{S}_1$. We then show how the simulator $\mathcal{S}_n$ works.

a) $\mathcal{S}_1$: The simulator $\mathcal{S}_1$ internally runs the adversary $\mathcal{A}_1$, emulates the token functionality $T^{\mathtt{HT}}$ for $\mathcal{A}_1$ for the session id sid and acts on the behalf of the parties $p_2, \ldots, p_n$. $\mathcal{S}_1$ executes the following steps.

Send $(e_2, f, \mathsf{vk}_2, \sigma_2, \mathsf{cert}_2), \ldots, (e_{n-1}, f, \mathsf{vk}_{n-1}, \sigma_{n-1}, \mathsf{cert}_{n-1})$ to $\mathcal{A}_1$.

If $\mathcal{A}_1$ sends $I = (\mathsf{Input}, \mathsf{sid}, x, f')$ to $T^{\mathtt{HT}}$ then do the following.

- If $\mathsf{ctr}_1^{\mathsf{sid}}$ is not defined then define it and set $\mathsf{ctr}_1^{\mathsf{sid}} \leftarrow 1$ otherwise output $\perp$ and stop.
- Output $(e_1, f, \mathsf{vk}_1, \sigma_1, \mathsf{cert}_1)$

If $\mathcal{A}_1$ sends $(e_1, f, \mathsf{vk}_1, \sigma_1, \mathsf{cert}_1)$ over broadcast then send (sid, x_1) to $\mathcal{F}$.

Upon receiving y_1 from $\mathcal{F}$ send $(\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ to $\mathcal{A}_1$.

Upon receiving $(\mathsf{Output}, \mathsf{sid}, z)$ from p_1, if $z = (\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ then send y_1 to $\mathcal{A}_1$, output $\perp$ otherwise.

b) $\mathcal{S}_n$: The simulator $\mathcal{S}_n$ internally runs the adversary $p_n^\star$, emulates the token functionality $T^{\mathtt{HT}}$ for $p_n^\star$ for the session id sid and acts on the behalf of the parties $p_1, \ldots, p_{n-1}$. $\mathcal{S}_n$ executes the following steps.

Send $((e_1, f, \mathsf{vk}_1, \sigma_1, \mathsf{cert}_1), \ldots, (e_{n-1}, f, \mathsf{vk}_{n-1}, \sigma_{n-1}, \mathsf{cert}_{n-1}))$ to $p_n^\star$.

If $I = (\mathsf{Input}, \mathsf{sid}, x_n, f', (e'_1, f_1, \mathsf{vk}'_1, \sigma'_1, \mathsf{cert}'_1), \ldots, (e'_{n-1}, f_{n-1}, \mathsf{vk}_{n-1}, \sigma'_{n-1}, \mathsf{cert}'_{n-1}))$ is received from $p_n^\star$, then check if for $j = 1, \ldots, n-1$: $e'_j = e_j$, $\sigma'_j = \sigma_j$, $\mathsf{vk}_j = \mathsf{vk}'_j$ and $\mathsf{cert}_i = \mathsf{cert}'_i$. If it is not, then output $\perp$ and stop, otherwise execute the following steps.

- If $\mathsf{ctr}_n^{\mathsf{sid}}$ is not defined then defined it and set $\mathsf{ctr}_n^{\mathsf{sid}} \leftarrow 1$ otherwise output $\perp$ and stop.
- Send x_n to $\mathcal{F}$ and upon receiving y_n from $\mathcal{F}$ send $(\overline{y}_1, \ldots, \overline{y}_{n-1}, f, \mathsf{vk}_n, \sigma, \mathsf{cert}_n)$ to $p_n^\star$.

The main differences from the real world are that the adversarial parties see dummy encryptions instead of the encryptions. We note that the simulation strongly relies on the fact that the adversaries cannot forge the signatures output of the hardware token functionalities. Indeed, by forging a signature an adversary could: 1) change the input of another party, or 2) use the inputs of the honest parties to evaluate a functions $f' \neq f$ or 3) evaluate the same function multiple times on the same honest parties inputs by changing the value sid thus completely breaking the security of the protocol.

In the case that all the parties are corrupted then we rely on the fact that $\mathcal{F}$ is CP-well-formed functionality and we allow the adversarial parties to communicate freely via $\mathcal{F}$. More precisely,

1) Whenever a message m is received on the ith adversarial interface, $\mathcal{F}$ outputs (i, m) to the first adversarial interface.
2) Whenever a message of the form (i, m) is received on the first adversarial interface, $\mathcal{F}$ outputs the message m to the ith adversarial interface.

□

B. Proof of Lemmata 1 and 2

To prove the security of Lemma 1 we rely on the fact that any execution of $\Pi^{\mathtt{HT\text{-}FBS}}$ can be seen as an execution of Π^{MPC} among honest parties. Indeed, in the case the hardware tokens are not corrupted the adversary has no control on the messages of Π^{MPC}, and he can only act as a proxy between the hardware tokens and the broadcast channel (since we assume that the adversary is non-aborting). This allows the CP simulators $\mathcal{S}_1, \ldots, \mathcal{S}_n$ to internally run the simulator Sim of Π^{MPC} (that exists by definition) for the case where no parties are corrupted. Hence, $\mathcal{S}_1, \ldots, \mathcal{S}_n$ can just run Sim using the same correlated randomness obtained from the trapdoor information $\widetilde{K}$, following the approach proposed in Sec V. To prove the security of Lemma 2 we can reduce the security of the entire protocol to the GUC security of Π^{MPC}. We note that in Lemma 2 the global setup is represented by $\overline{T}^{\mathtt{HT\text{-}FBS}}$, which captures the scenario where the hardware tokens are compromised. We now provide more formal arguments.

1) Proof of Lemma 1.: We start the proof by assuming that at least one party is honest. In order to prove this part of the theorem we need to show a collection of efficiently computable transformations $\mathsf{Sim} = \{\mathsf{Sim}_1, \ldots, \mathsf{Sim}_n\}$ mapping ITMs to ITMs such that for every set of adversaries $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, and every PPT environment $\mathcal{Z}$ the following holds:

$$\text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi^{\mathtt{HT\text{-}FBS}},\mathcal{A},\mathcal{Z}} \approx \text{CP-EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\phi,\mathsf{Sim}(\mathcal{A}),\mathcal{Z}}$$

By assumption on the security of the MPC protocol, we know that $\forall \mathcal{A}\ \exists \mathsf{S}\ \forall \mathcal{Z}$ such that $\text{EXEC}^{\bar{\mathcal{G}},\mathcal{B}}_{\Pi^{\mathsf{MPC}},\mathcal{A},\mathcal{Z}} \approx \text{EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\mathsf{S},\mathcal{Z}}$.

We consider now S for the case where there are no corrupted parties and describe how $\mathcal{S}_i = \mathsf{Sim}(\mathcal{A})_i$ works for $i = 1, \ldots n$. Without loss of generality we formally describe how the simulator $\mathcal{S}_1$ works since the other simulators follow exactly the same strategy.

The simulator $\mathcal{S}_1$ queries $(\mathsf{GetTrapdoor}, \mathsf{sid})$ $\mathcal{F}^\star$ with the command $(\mathsf{GetTrapdoor}, \mathsf{sid})$. $\mathcal{F}^\star$ checks that sid is equal to its session id. If it is, then $\mathcal{F}^\star$ sends $(\mathsf{Trapdoor}, \mathsf{sid})$ to $T^{\mathtt{HT\text{-}FBS}}$.

$T^{\mathtt{HT\text{-}FBS}}$ sends $\{\mathsf{vk}_i, \mathsf{cert}_i\}_{i\in[n]}, \{\mathtt{msg}_j^\ell, \sigma_j^\ell\}_{j\in[n],\ell\in[m]}$ information to $\mathcal{F}^\star$ which forwards them to the simulator $\mathcal{S}_1$. Then $\mathcal{S}_1$ executes the following steps.

- Send $(\mathtt{msg}_2^\ell, \mathsf{vk}_2, \sigma_2^\ell, \mathsf{cert}_2), \ldots, (\mathtt{msg}_n^\ell, \mathsf{vk}_n, \sigma_n^\ell, \mathsf{cert}_n)$ to $\mathcal{A}_1$.
- If $I = (\mathsf{Input}, \mathsf{sid}, x_1, R_1, \Pi^{\mathsf{MPC}})$ is received from $\mathcal{A}_1$ then do the following.
  - If $\mathsf{ctr}^{\mathsf{sid}} = 0$ then $\mathsf{ctr}^{\mathsf{sid}'} \leftarrow 1$ otherwise output $\perp$ and stop.
  - Set $\mathtt{x} \leftarrow x_1$, $\mathtt{l} \leftarrow 1$ and send $(\mathtt{msg}_1^1, \Pi^{\mathsf{MPC}}, \mathsf{vk}_1, \sigma_1^1, \mathsf{cert}_1)$ to $\mathcal{A}_1$.
- If $I = (\mathsf{NextMsg}, \mathsf{sid}, \{\mathtt{msg}_i^{\ell'}, \sigma_i^{\ell'}, \mathsf{vk}_i{}', \mathsf{cert}_i{}'\}_{i\in[n]})$ is received then do the following.

- If $\ell \neq \mathtt{l}$ or $\mathtt{l} > m$ output $\perp$ and stop, otherwise continue with the following steps.
- For all $j \in [n]$ if $\mathtt{msg}_j^{\ell} \neq \mathtt{msg}_j^{\ell'}$ or $\sigma_j^{\ell} \neq \sigma_j^{\ell'}$ or $\mathsf{cert}_i{}' \neq \mathsf{cert}_i$ or $\mathsf{vk}_i{}' \neq \mathsf{vk}_i$ then output $(\perp, p_i)$ and stop.
- Set $\mathtt{l} \leftarrow \mathtt{l} + 1$.
- If $\ell \leq m$ then send $(\mathtt{msg}_1^{\mathtt{l}}, \mathsf{sid}, \Pi^{\mathsf{MPC}}, \sigma_1^{\mathtt{l}})$ to $\mathcal{A}_1$.
- If $\ell = m + 1$ then send x to $\mathcal{F}$. Upon receiving y_1 from $\mathcal{F}$ send y_1 to $\mathcal{A}_1$.

The proof relies on the observation that (unless the adversary breaks the security of the strong signature scheme), then the adversary just acts as an observer on the channel. That is, the corrupted party $\mathcal{A}_1$ can only inspect the messages generated by $T_1^{\text{HT-FBS}}$ and the messages received on the channel which are honestly generated using $T^{\text{HT-FBS}}$.[10] For this reason $\mathcal{S}_1, \ldots, \mathcal{S}_n$ can run a simulator S of Π^{MPC} that works when there are no corrupted parties. We recall that the behavior of the corrupted parties cannot influence the output of S as long as the signature scheme is secure.

2) Proof of Lemma 2.: This proof follows immediately from the GUC security of Π^{MPC}. Indeed, we note that the token functionalities simply run Π^{MPC} even in the case that a unsigned message is received (or a message is not received at all). This also enables identifiable abort (unanimous abort) if Π^{MPC} enables it.

C. Proof Sketch of Theorem 2

The only way that an adversary has to misbehave is by sending an invalid signature over the channel. This behavior can be detected by any party that has the verification keys of the token functionalities. Moreover, the first party that sends a unsigned message is identified as corrupted. In all the schemes that we have described in this section, the broadcast channel is used exclusively to run the protocol. Hence, once the protocol has been executed the broadcast channel is closed (no messages can be send on that channel). We note that this is not a limitation since the notion of CP tolerates composability. Hence, multiple broadcast channels might be available at the same time.

D. Proof of Theorem 3

Proof. To prove this theorem we rely on the observation that Π^{HT} ($\Pi^{\text{HT-FBS}}$) is collusion-preserving for $\mathcal{F}^f$ as long as nobody aborts. That is, the only way to do subliminal communication in Π^{HT} (and in $\Pi^{\text{HT-FBS}}$) is by sending a message which is incompatible with the protocol description thus causing the protocol to abort. Given the way we have set the payoffs, it is always inconvenient for the adversary to trigger the collusion event $E_{\mathtt{collude}}$, as it causes the abort event $E_{\mathtt{abort}}$. We observe that in the case where we also consider the utility of the designer we can prove that our protocol is CPIC (according to Def. 4) in the case there the utilities of the designer are symmetric to the utility of the adversary.

[10] If that is not the case then either the protocol would abort, or a reduction to the security of the signature scheme can be done.

In the case that all the parties are malicious then we rely on the fact that $\langle \mathcal{F}^f \rangle$ is CP-well-formed functionality and we allow the adversarial parties to communicate freely via $\langle \mathcal{F}^f \rangle$ following the same approach proposed in the proof of Theorem 1.

In the case that at least one party is honest then we need to show collection of efficiently computable transformations $\mathsf{Sim} = \{\mathsf{Sim}_1, \ldots, \mathsf{Sim}_n\}$ mapping ITMs to ITMs such that for every set of adversaries $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, and every PPT environment $\mathcal{Z}$ the following holds:

$$\text{CP-EXEC}^{\bar{\mathcal{G}}, \mathcal{R}}_{\Pi^{\text{HT}}, \mathcal{A}, \mathcal{Z}} \approx \text{CP-EXEC}^{\bar{\mathcal{G}}, \langle \mathcal{F}^f \rangle}_{\phi, \mathsf{Sim}(\mathcal{A}), \mathcal{Z}}$$

The simulators are equal to the simulators showed in the proof of Theorem 1 except for the following details. If $\mathcal{S}_i$ receives an invalid message m from $p_i^\star$ (i.e. a message that would yield an honest party to abort) then $\mathcal{S}_i$ sends $(\mathtt{COLLUDE}, \mathsf{sid}, m)$ to $\langle \mathcal{F}^f \rangle$ and stops. This situation captures the ability of the corrupted parties to interact with each others at the price of being detected by the honest parties. Indeed, we recall that Π^{HT} enjoys the identifiable abort property. The payoff we have defined ensures that an adversary is never incentivized to abort (i.e. break the CP property). Given that we have proved in Theorem 1 that the if no party aborts then Π^{HT} is collusion-preserving then we can claim that Π^{HT} is $CPAP$. The same arguments can be applied to $\Pi^{\text{HT-FBS}}$.

□

E. Proof of Theorem 4

Proof Sketch of Theorem 4. To prove Π_{pen} realizes the CP-functionality $\mathcal{F}$, $\mathcal{S}_i$ (the simulator for party p_i) communicates with $\mathcal{F}_{\text{pen}}$ to determine the status of the deposited collateral. To simulate Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$) it follows the simulator from the proof of Theorem 1 (resp. Lemma 1). To prove CPAP, we show that Π_{pen} realizes $\langle \mathcal{F} \rangle$ by letting $\mathcal{S}_i$ in addition use the $\mathtt{COLLUDE}$ command in $\langle \mathcal{F} \rangle$ in case the adversary misbehaves during Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$). Since in Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$) collusion always follows an abort, CPAP of Π_{pen} follows from setting the deposit d so that the cost of abort (losing td collateral) is higher than gains from collusion $\gamma_{\mathtt{collude}}$—that is, the adversary cannot gain more utility via triggering weaknesses in $\langle \mathcal{F} \rangle$. Lastly, to prove CPIC, we show that given a passive adversarial strategy $\mathcal{A}$ that just follows Π_{pen}, the designer has no incentive to deviate from choosing Π_{pen}. We show this by observing that $u_{\mathtt{D}}(\Pi_{\text{pen}}, \mathcal{A}) = 0$, which is the maximum utility of the designer (as parties can only reclaim collateral which they have already deposited). □

Full proof of Theorem 4. First, we prove that Π_{pen} realizes the CP-functionality $\mathcal{F}$ as long as there is no abort. To emulate Π_{pen}, the simulator $\mathcal{S}_i$ forwards $(\mathtt{DEPOSIT})$ and $(\mathtt{RECLAIM})$ to $\mathcal{F}_{\text{pen}}$ on behalf of corrupt party p_i and returns whatever $\mathcal{F}_{\text{pen}}$ does to p_i. During Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$), $\mathcal{S}_i$ follows the simulator for p_i in Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$). By Theorem 1 (resp. Lemma 1) which proved that Π^{HT} (resp. $\Pi^{\text{HT-FBS}}$) realize $\mathcal{F}$ when there is no abort, $\mathcal{S}_i$ emulates party p_i in Π_{pen} when there is no abort.

Second, we prove that Π_{pen} is CPAP-secure. To do so, we must show that (1) Π_{pen} realizes $\langle\mathcal{F}\rangle$, and (2) the adversary has no incentive to deviate from Π_{pen} in such a way that forces the simulators to use weaknesses in $\langle\mathcal{F}\rangle$ (collude/abort) in order to simulate. To show (1), we construct a simulator $\mathcal{S}_i$ for each corrupt party p_i. First $\mathcal{S}_i$ sends $(\texttt{DEPOSIT})$ on behalf of p_i (or not, if p_i chooses not to). Then, $\mathcal{S}_i$ is made aware of whether all parties have sent $(\texttt{DEPOSIT})$ to $\mathcal{F}_{\text{pen}}$, when $\mathcal{F}_{\text{pen}}$ returns the message $(\texttt{DEPOSIT}, 1)$ or $(\texttt{DEPOSIT}, 0)$. If $(\texttt{DEPOSIT}, 0)$ is returned, the protocol stops. To simulate Step 2 of Π_{pen}, we simply do as the simulator for $\Pi^{\texttt{HT}}$ (resp. $\Pi^{\texttt{HT-FBS}}$) would, except if it sees an invalid message m, then $\mathcal{S}_i$ sends $(\texttt{COLLUDE}, \mathsf{sid}, m)$ to $\langle\mathcal{F}\rangle$. Finally, $\mathcal{S}_i$ sends $(\texttt{RECLAIM})$ to $\mathcal{F}_{\text{pen}}$ if p_i does, and forwards $(\texttt{RECLAIM}, 1)$ or $(\texttt{RECLAIM}, 0)$ from $\mathcal{F}_{\text{pen}}$.

To prove (2), we show that the adversary has no incentive to deviate from Π_{pen} at all. Now, from our utility function $u_{\mathtt{A}}$, we see that the only way the adversary can possibly gain utility by deviating is by triggering the event $E_{\texttt{collude}}$ (since in the protocol he can only reclaim collateral he has deposited, collusion is the only possible way to gain positive utility). We observe the only time we may need to use $\texttt{COLLUDE}$ is during simulating Step 2 (running $\Pi^{\texttt{HT}}$ or $\Pi^{\texttt{HT-FBS}}$). If collusion occurs during Step 2, at least one honest party will refuse to send $(\texttt{RECLAIM})$. Thus, the attacker achieves payoff of $-td + \gamma_{\texttt{collude}}$ in this scenario (where t is the number of corrupt parties). This means deviation is not profitable and we obtain CPAP, when we set deposit d such that $td > \gamma_{\texttt{collude}}$.

Finally, to achieve CPIC, we must find an attacker $\mathtt{A}$ such that $(\Pi_{\text{pen}}, \mathtt{A})$ is a $\nu(\lambda)$-subgame perfect equilibrium where the designer can arbitrarily choose another protocol in $\mathbf{\Pi}$. Denote Π_{pen} as Π for ease of reading. Let $\mathtt{A}$ be the passive attacker who simply follows Π_{pen}. We now prove the two conditions of subgame perfect equilibrium, which are: (1) for any $\Pi' \in \mathbf{\Pi}$, $u_{\mathtt{D}}(\Pi', \mathtt{A}(\Pi')) \leq u_{\mathtt{D}}(\Pi, \mathtt{A}(\Pi)) + \epsilon$, and (2) for any $\mathtt{A}' \in \texttt{ITM}$, $u_{\mathtt{A}}(\Pi, \mathtt{A}'(\Pi)) \leq u_{\mathtt{A}}(\Pi, \mathtt{A}(\Pi)) + \epsilon$. We have already shown (2) in the CPAP proof above. For (1), we observe that the maximum utility possible for the designer is zero, as $\mathcal{F}_{\text{pen}}$ only allows the parties to reclaim what they have deposited. We also observe that except for negligible probably (that the adversary breaks e.g. the signature scheme), for the passive attacker $\mathtt{A}$, the designer achieves zero utility if he chooses $\Pi = \Pi_{\text{pen}}$ (everyone deposits collateral, follows the protocol, and then reclaims), thus $u_{\mathtt{D}}(\Pi, \mathtt{A}(\Pi))$ achieves maximum utility (minus negligible utility loss from the negligible probability of collusion via the adversary breaking e.g. signatures). This implies (1).

□

APPENDIX C
IMPOSSIBILITY RESULTS IN THE MEDIATED MODEL

In Lemma 2 and Theorem 2, we showed that even if our assumptions on honest hardware tokens are broken, we can still achieve fallback security with identifiable and unanimous abort.

On the other hand, desirable properties—robustness, fairness, and identifiable abort—are impossible in the mediated model when the mediator may be corrupt. This implies undesirable scenarios—a game of poker implemented in this model can be aborted at any time (for example, when a corrupt party sees he is losing), without honest players being able to accuse the cheater.

We recall that in the mediated model, every party communicates only to a central node called the mediator, in a star-topology network. Intuitively, these impossibilities stem from the mediator being able to cut off communication between itself and any party, at any time. Robustness is simple to show to be impossible, since the mediator can simply stop all communication. To break fairness, the mediator can end communication in the protocol after one party receives his output and before another party receives his. This strategy can be used also to break unanimous abort and allow one honest party to receive his output while others do not. Lastly, identifiability is not possible as a corrupt mediator can "frame" an honest party as having aborted, by simply ignoring messages from this party. Since other parties only communicate through the mediator, they cannot identify whether the party or the mediator has misbehaved.

We state and prove these impossibilities formally in the theorems below.

Theorem 6. *Robustness is not possible when the mediator may be corrupted. This holds even when all other parties are honest.*

Proof. The mediator simply does nothing, and thus nothing can be computed. □

Theorem 7. *Consider a protocol in which each party interacts with the mediator a number of times that is polynomial in the security parameter. Then fairness and unanimous abort are not possible when the mediator may be corrupted, unless the output can be computed without the parties' inputs. This holds even if all other parties are honest.*

Proof. Consider in the mediated model three parties: (honest) parties Alice (A) and Bob (B), and the corrupt mediator (M) who has no inputs and some constant as output (e.g., 1). Since the mediator controls when messages are sent during the protocol, we consider the following protocol format without loss of generality: In odd-numbered interactions, A interacts with M; in even-numbered interactions, B interacts with M.

Suppose M does the following: M guesses an interaction number in which only one of A or B has the output (this is always possible since A and B never receive messages at the same time and in the mediated model, A and B cannot communicate except through M), and aborts at the beginning of this interaction, depriving the other party of the output. Since there is a polynomial number of choices, M succeeds with non-negligible probability.

Formally: Suppose the final interaction in the protocol is between A and M. Assuming that fairness is possible, we show

that if this protocol has fairness, then A and B can compute their outputs without sending any messages.

Since the final interaction does not involve B, B must have known the output prior to the final interaction. Suppose the corrupt M chooses to abort in the final interaction (Since the number of rounds is polynomial, he chooses to abort at this round with non-negligible probability). By fairness (since B knows the output without the final message), A must also be able to learn the output without the final interaction. This means both A and B learn the output without the final interaction.

Similarly, the second-to-last interaction (between B and M) does not involve A, so A must have known the output prior to this interaction. By fairness, if M chooses to abort in the second-to-last interaction, B must also be able to learn the output without this interaction. Thus both A and B learn the output without the second-to-last round.

Continue the argument for the number of interactions, and A and B both learn the output without sending any messages. □

The proof does not work when the number of interactions is superpolynomial. This is because the mediator will only guess correctly with negligible probability, which interaction to abort at (we allow negligible failure of the fairness property).

Theorem 8. *Identifiability is not possible in the mediated model when the mediator may be corrupted, regardless of the number of honest parties.*

Proof. We show that it is impossible to distinguish between the case when 1) the mediator is corrupt, and 2) a party A aborts at round r. Suppose M is corrupt and does the following: M follows the protocol until round $r-1$. At rounds $\geq r$, M simply ignores (does not send nor receive from) A, and does whatever it is supposed to do in the protocol if A aborts. Since all parties communicate only via M, the messages they receive in this scenario are exactly the same as if the mediator were honest, but a corrupt party A has followed the protocol before round r and stops sending messages after round r. Thus, they cannot correctly identify whether M or A is corrupt in the case of an abort. □

APPENDIX D
MOTIVATION FOR USING PENALIZATION TO DISINCENTIVIZE ABORTS

We observe that the notions of collusion-preservation (and so collusion-freeness) do not capture all the attacks that send subliminal message using the ability to abort. Indeed and adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ could adapt the following strategy in a game of poker. $\mathcal{A}_1$ aborts any time that he holds an Three of Spades in his hand. Clearly, even if $\mathcal{A}_1$ and $\mathcal{A}_2$ are isolated, when protocol does not abort, $\mathcal{A}_2$ knows some information about $\mathcal{A}_1$'s cards that the honest parties do not. This attack becomes more interesting in the case of reactive functionalities, where the parties can get intermediate outputs. As a concrete example of a reactive functionality, we can think to use a CP protocol to shuffle a deck of cards and let the parties play poker. A strategy could be to make corrupt party $\mathcal{A}_1$ abort in the first hand if he does not have an Ace, and abort in the second hand if he does not have a King, and continues the game otherwise. It should be easy to see that with high probability $\mathcal{A}_1$ will be able to communicate more than one bit of information to $\mathcal{A}_2$. We note that this issue arises only because the CP definition allows the parties to abort, which is in general an unavoidable attack. Thus, in this work we choose to disincentivize aborts.

A. Open Questions: An alternative penalization functionality

The penalization solution Π_{pen} in Sec. VI-C, while simple and satisfying CPAP and CPIC, requires at least one honest party to penalize the misbehaving party even at the cost of his own collateral being lost. In fact, if a corrupt party aborts during protocol execution, then honest parties can increase their utility by ignoring the abort and still allow everyone to reclaim their collateral. In other words, a strategy profile $(\Pi, \mathcal{A})$, where Π is the same as Π_{pen} *except* it allows parties to always reclaim their collateral, and $\mathcal{A}$ is an adversarial strategy that makes a corrupt party collude and abort, is also an equilibrium solution.

Below, we suggest below an alternative penalization functionality $\mathcal{F}_{\text{penalize}}$ (Fig. 7) that only punishes misbehaving parties and thus eliminates this unwanted equilibrium. It is clear to see that if only misbehaving parties are punished, honest parties will have no incentive to deviate from the penalization scheme, as they will always be able to reclaim their deposit. In addition, this solution does not require one party to be honest, as the functionality $\mathcal{F}_{\text{penalize}}$ is responsible for verifying whether the protocol aborted and who is responsible. However, this functionality is more complex and is more difficult to realize than $\mathcal{F}_{\text{pen}}$, as discussed below.

Informally, a protocol Π_{penalize} running $\mathcal{F}_{\text{penalize}}$ disincentivizes misbehavior by allowing parties to run a protocol only if all parties have deposited a collateral. Throughout the protocol, it keeps track of `PROTOCOL` messages submitted by the parties. A party can reclaim his collateral if he has behaved correctly throughout protocol execution (according to the `PROTOCOL` messages). Otherwise, if the protocol aborts, by using the publicly identifiable abort property, the functionality ensures that only identified corrupt parties lose their collateral.

a) Implementing $\mathcal{F}_{penalize}$ on the blockchain: some caveats.: A natural question is whether $\mathcal{F}_{\text{penalize}}$ can be implemented on the blockchain, for example following the approach of [39]. There, parties make deposits on the blockchain before running an MPC protocol, and can only reclaim their deposits if they submit correct messages to the blockchain at each round. In particular, if the MPC protocol has constant number of rounds, [39] only requires a constant number of ledger rounds.

However, a concerning caveat to implementing $\mathcal{F}_{\text{penalize}}$ in the blockchain is that Π^{HT} and $\Pi^{\text{HT-FBS}}$ now run using the blockchain, a (usually) non-CP protocol, as its communication channel. In particular, this means that *within the protocols* Π^{HT} and $\Pi^{\text{HT-FBS}}$, corrupt parties might trivially collude via its

communication resource. We contrast this with the situation of Π_{pen}, our penalization protocol in Sec. VI-C. In Π_{pen}, parties run Π^{HT} and $\Pi^{\text{HT-FBS}}$ via broadcast, and do not require the blockchain while executing these protocols. Even if we assume that the blockchain is accessible to some parties at some point during this execution, we can handle the access as an external communication resource. Collusion-preservation ensures that even given some external communication, the correlation between corrupt parties in $\Pi^{\text{HT}}/\Pi^{\text{HT-FBS}}$ does not increase substantially. For example, a corrupt party cannot simply post some secret key on the blockchain that allows him to arbitrarily subliminally communicate with others. In addition to this caveat, another concern is the difficulty keeping track of protocol messages on the blockchain. In general, one cannot guarantee that transactions sent at some round r will be in the blockchain ledger at round r, as is the case in the analyses of, e.g. [7]. In contrast, in $\mathcal{F}_{\text{pen}}$ (Sec. VI-C) protocol messages of Π^{HT} or $\Pi^{\text{HT-FBS}}$ use the broadcast channel instead of the blockchain, and thus are not subject to idiosyncrasies of the blockchain.

B. Protocol $\Pi_{penalize}$ for penalization

We describe below Π_{penalize}, which uses the penalization functionality $\mathcal{F}_{\text{penalize}}$ (Fig. 7). The only differences between Π_{penalize} and Π_{pen} are that Π_{penalize} runs the entire protocol using $\mathcal{F}_{\text{penalize}}$ to communicate, and it penalizes only corrupt parties for misbehavior. That is, honest parties never lose their collateral when playing Π_{penalize}.

Protocol Π_{penalize}:

- Each party sends $(\texttt{DEPOSIT})$ to $\mathcal{F}_{\text{penalize}}$, to deposit a collateral of amount d. If the functionality returns $(\texttt{DEPOSIT}, 1)$, proceed. Otherwise, stop.
- Run Π^{HT} or $\Pi^{\text{HT-FBS}}$ on $\mathcal{F}_{\text{penalize}}$, by sending/receiving each protocol message m as $(\texttt{PROTOCOL}, m)$ to/from $\mathcal{F}_{\text{penalize}}$.
- All parties send $(\texttt{RECLAIM})$ to $\mathcal{F}_{\text{penalize}}$.

We can thus prove a theorem about CP, CPAP, and CPIC for Π_{penalize}, analogous to Theorem 4, using the same utility functions $u_{\texttt{A}}$ and $u_{\texttt{D}}$ defined in Sec. VI-C.

Theorem 9. *The $\{\bar{\mathcal{G}}\}$-exclusive protocol $\Pi_{penalize}$, where $\bar{\mathcal{G}} = T^{\text{HT}}, \mathcal{F}_{penalize}$ (resp. $\bar{\mathcal{G}} = T^{\text{HT-FBS}}, \mathcal{F}_{penalize}$) satisfies the same theorem statements as Theorem 4, except assuming $d > \gamma_{\texttt{collude}}$.*

Proof. The proof is exactly the same as Theorem 4, except instead of running on the broadcast channel the protocol runs entirely on $\mathcal{F}_{\text{penalize}}$ (accessible to the simulator, which makes the simulator's job easier than having to simulate broadcast messages in Π_{pen}). For CPAP, an attacker can now trigger $E_{\texttt{collude}}$ by only having *one* corrupt party abort and lose his collateral. That is, we can have CPAP only if $d > \gamma_{\texttt{collude}}$ (i.e., the loss of one party's collateral is more than the gain from colluding). As for the designer $\texttt{D}$, same as in Π_{pen} the maximum designer utility is still zero, and thus the CPIC proof still holds. □

The functionality is parameterized by a deposit amount d, a tuple of initial balances of each party, $\mathsf{Balance} = (b_1, \ldots, b_n)$, and a protocol Π which is either Π^{HT} or $\Pi^{\text{HT-FBS}}$. It keeps track of PState, initialized as $\perp$, which keeps track of all protocol messages $(\texttt{PROTOCOL}, \cdots)$.

- Upon receiving $(\texttt{DEPOSIT})$ from party p_i: If $b_i \geq d$:
  In this round, if all $p_j \in \mathcal{P}$ have sent $(\texttt{DEPOSIT})$, and $b_j \geq d$, then continue. Otherwise, send $(\texttt{DEPOSIT}, 0)$ to each party $p_i \in \mathcal{P}$ and stop.
  Deposit collateral for each party: for each party $p_i \in \mathcal{P}$: $b_i \leftarrow b_i - d$
  Send the message $(\texttt{DEPOSIT}, 1)$ to each party $p_i \in \mathcal{P}$.
- For each round of Π, upon receiving $(\texttt{PROTOCOL}, m)$ from party p_i:
  Verify this message via Π and PState. If Π aborts, stop processing $\texttt{PROTOCOL}$ messages.
  Send $(\texttt{PROTOCOL}, m)$ to all parties and store it in PState.
- Upon receiving $(\texttt{RECLAIM})$ from party p_i, check PState to see if p_i has misbehaved[a]:
  If p_i behaved correctly, return the deposit to p_i: $b_i \leftarrow b_i + d$ and send $(\texttt{RECLAIM}, 1)$ to p_i.
  Otherwise, send $(\texttt{RECLAIM}, 0)$ to p_i.

[a] This can be checked as Π has publicly identifiable abort

Fig. 7. $\mathcal{F}_{\text{penalize}}$

APPENDIX E
A NOTE ON GAMES WITH SHORT STRATEGY DESCRIPTION

In [45] it is observed that it would be trivial to avoid collusion for reactive functionalities (e.g. games) in the following model: The players commit to their strategy (when the strategy can be described in a short way) and then run a secure function evaluation on these commitments, aborting when parties deviate from their committed strategies. Indeed, once the strategy of a game is committed in a collusion-preserving manner, if the parties are not allowed to abort, then any subliminal communication between the malicious parties cannot harm the outcome of the game. The situation is more complicated in the case that the malicious parties are allowed to abort since such a behaviour could cause the game to stop. However, aborts can be disincentivized by a penalization scheme such as the one proposed in Section VI-C.

APPENDIX F
THE FUNCTIONALITY $\mathcal{F}_{\text{PEN}}$

We present our penalization functionality in Figure 8.

APPENDIX G
A NOTE ON CORRELATED EQUILIBRIA

In strategic games, players can achieve correlated equilibrium by observing the same public signal. In our construction,

The functionality is parameterized by a deposit amount d, and a tuple of initial balances of each party, $\mathsf{Balance} = (b_1, \ldots, b_n)$.

- Upon receiving $(\texttt{DEPOSIT})$ from party p_i: If $b_i \geq d$:
  In this round, if all $p_j \in \mathcal{P}$ have sent $(\texttt{DEPOSIT})$, and $b_j \geq d$, then continue. Otherwise, send $(\texttt{DEPOSIT}, 0)$ to each party $p_i \in \mathcal{P}$ and stop.
  Deposit collateral for each party: for each party $p_i \in \mathcal{P}$: $b_i \leftarrow b_i - d$
  Send the message $(\texttt{DEPOSIT}, 1)$ to each party $p_i \in \mathcal{P}$.
- Upon receiving $(\texttt{RECLAIM})$ from party $p_i \in \mathcal{P}$:
  In this round, if received $(\texttt{RECLAIM})$ from all $p_i \in \mathcal{P}$, continue to return deposits to all parties. Otherwise, return $(\texttt{RECLAIM}, 0)$ and stop.
  For each $p_i \in \mathcal{P}$, return the deposit to p_i: $b_i \leftarrow b_i + d$ and send the message $(\texttt{RECLAIM}, 1)$.

Fig. 8. $\mathcal{F}_{\text{pen}}$ allows the parties to deposit an amount d, which can be withdrawn only if all the parties registered to the functionality send a command RECLAIM.

this public signal can be the setup (e.g. public keys) or the broadcasted protocol messages. Similarly, the CP construction in the mediated model by Alwen et al. introduces a public signal via a GUC-complete setup (e.g., ACRS); see Theorem 5.1 of [2]. Indeed, this introduces additional correlated equilibria to the ideal game; however, collusion preservation ensures that no meaningful information (about the input or output) can be communicated using the additional correlation.

APPENDIX H
FORMAL DEFINITIONS

1) Hardware tokens and setup assumptions.: Our protocols assume that each party has access to a stateful hardware token which might perform fresh encryptions and signatures with respect to some hidden keys. Moreover, an adversary that has physical access to the HT can only query it and get the result back, without having the possibility to inspect the intermediate steps of the computation or tamper with the token. A similar assumption is used in [3] to generate randomness hidden from the adversary, in order to obtain receipt-freeness. The work of [50] provides a formal and composable abstraction of the Intel SGX hardware token, which has all the features we have just described. In this work, following [3], we abstract the hardware token by means of an ideal functionality that has a secret state (the secret keys), a public state (the public key and the signature verification key) and code. All the outputs of the HT are signed together with the code that has been run to generate the output. This process is called *attestation* and assures parties holding the verification key of the HT that the output has been generated following a specified code. We note that by assuming the existence of hardware tokens we are also implicitly assuming a setup assumption that is stronger than a public key infrastructure (PKI). Indeed, we require that all parties running our protocol know each other hardware token public keys and that those public keys are generated honestly. However, we note this approach is still meaningful in the context of collusion-freeness/collusion-preservation since the hardware token (and the public keys) are generated before inputs are given to the parties and even before the function being computed is agreed upon.

In this paper we also consider the case where hardware tokens are corrupted. That is, the adversary can reprogram all malicious parties' tokens, and can learn the secret keys of every (both malicious and honest) party's token. This corruption model captures the real world scenario where an adversarial party can break the security of his own token, but could only obtain the secret key of tokens that he has no physical access to.

2) Negligible functions.: We say a function ν is *negligible* if for every positive integer c there is an integer N_c such that for all $x > N_c$, $|\nu(x)| < 1/x^c$. We say $a \overset{\text{negl}}{\leq} b$ for real values a, b to mean that $a \leq b + \nu(\lambda)$ for negligible function ν.

3) Secure function evaluation.: We consider a protocol Π^{MPC} that securely (GUC-)realises any efficient functionality $\mathcal{F}$. We additionally assume that Π^{MPC} can be run over a broadcast channel.

Following [10] we consider MPC protocols where at each round ℓ, each party P_i broadcasts a message $\mathtt{msg}_i^\ell$ to all the other parties simultaneously.

Definition 5 (MPC protocol)**.** *Let n be a positive integer, m a polynomial in the security parameter, and $\mathcal{F}$ an n party-functionality. An m-round, n-party MPC protocol Π^{MPC} for $\mathcal{F}$ can be described as a tuple of deterministic polynomial-time algorithms $\Pi^{\mathsf{MPC}} = (\mathsf{Next}_1, \ldots, \mathsf{Next}_n)$.*

Next message: $\mathtt{msg}_i^\ell \leftarrow \mathsf{Next}_i(1^\lambda, x_i, \rho_i^\ell, \overline{\mathtt{msg}}^{<\ell})$ is the message broadcasted by party $p_i \in \mathcal{P}$ in round $\ell \in [m]$, on input $x_i \in \{0,1\}^\lambda$, on random tape $\rho_i^\ell \xleftarrow{\$} \{0,1\}^\lambda$, after receiving the messages $\overline{\mathtt{msg}}^{<\ell} = \{\mathtt{msg}_j^{\ell'}\}_{j\in[n],\ell'<\ell}$, where $\mathtt{msg}_j^{\ell'}$ is the message broadcasted by party p_j on round $\ell' \in [\ell - 1]$. When $\overline{\mathtt{msg}}^{<m+1} = \{\mathtt{msg}_j^{\ell'}\}_{j\in[n],\ell'<m+1}$ then $y_i \leftarrow \mathsf{Next}_i(1^\lambda, x_i, \rho_i^{m+1}, \overline{\mathtt{msg}}^{<m+1})$ where y_i denotes the output of the party p_i.

In this paper we make use of a protocol Π^{MPC} that GUC-realizes a functionality $\mathcal{F}$ where the resource $\mathcal{R}$ is a broadcast channel. More formally, we require that $\forall\mathcal{A}\ \exists\mathsf{Sim}\ \forall\mathcal{Z}$ such that $\text{EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi^{\mathsf{MPC}},\mathcal{A},\mathcal{Z}} \approx \text{EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\mathsf{Sim},\mathcal{Z}}$. We refer the reader to App. I for the formal definition of GUC security.

Definition 6 (Strong unforgeable signature scheme)**.** *A triple of* PPT *algorithms $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{Ver})$ is called a* signature scheme *if it satisfies the following properties.*

Completeness: *For every pair $(s, v) \xleftarrow{\$} \mathsf{Kgen}(1^\lambda)$, and every $m \in \{0,1\}^\lambda$, we have that $\Pr[\mathsf{Ver}(v, m, \mathsf{Sign}(s, m)) = 0] = 1 - \nu(\lambda)$.*

Consistency (non-repudiation): *For any m, the probability that $\mathsf{Kgen}(1^\lambda)$ generates (s, v) and $\mathsf{Ver}(v, m, \sigma)$ generates two different outputs in two independent invocations is smaller than $\nu(\lambda)$.*

(Strong) Unforgeability: *For every* PPT $\mathcal{A}$*, there exists a negligible function* ν*, such that for all auxiliary input* $z \in \{0,1\}^\star$ *it holds that:*

$$\Pr[(s,v) \xleftarrow{\$} \mathsf{Kgen}(1^\lambda); (m,\sigma) \xleftarrow{\$} \mathcal{A}^{\mathsf{Sign}(s,\cdot)}(z,v) \wedge \mathsf{Ver}(v,m,\sigma) = 1 \wedge (m,\sigma) \notin Q] < \nu(\lambda)$$

where Q *denotes the set of the couples message-signature* $\{(m_i,\sigma_i)\}_{i\in\lambda}$ *where* m_i *is requested by* $\mathcal{A}$ *to the oracle* $\mathsf{Sign}(s,\cdot)$ *which returns* $\sigma_i \xleftarrow{\$} \mathsf{Sign}(s,m_i)$ *for all* $i \in \{1,\ldots,\lambda\}$.

In this paper we also make use of the UC-signature functionality proposed in [16] that we denote with $\mathcal{F}_{\mathtt{SIGN}}$. We also use the the fact that a scheme $\Sigma = (\mathsf{Gen},\mathsf{Sign},\mathsf{Ver})$ that satisfies Def. 6 can be turned into a scheme that UC-realized the functionality $\mathcal{F}_{\mathtt{SIGN}}$ [16, Thm 2]. We assume familiarity with $\mathcal{F}_{\mathtt{SIGN}}$, and for more discussion on this functionality we refer the reader to Sec. I-A Fig. 9.

Definition 7 (CPA-secure Symmetric Encryption Scheme (from notes of [30], Definition 6.8))**.** *A triple of* PPT *algorithms* $\mathsf{SE} = (\mathsf{Gen},\mathsf{Enc},\mathsf{Dec})$ *is called a* chosen-plaintext-attack-secure symmetric encryption scheme *if it satisfies the following properties.*

Completeness*: For every secret key* $s \xleftarrow{\$} \mathsf{Gen}(1^\lambda)$*, and every* $m \in \{0,1\}^\lambda$*, we have that* $\Pr[\mathsf{Dec}(s,\mathsf{Enc}(s,m)) = m] = 1$.

CPA-Security*: Let the left-or-right encryption oracle be as follows, where* $b \in \{0,1\}$*,* $m_0, m_1 \in \{0,1\}^\lambda$*:*

Oracle $\mathsf{Enc}(\mathsf{LR}(m_0,m_1,b))$*:*
if $|m_0| \neq |m_1|$ then return $\perp$
$c \xleftarrow{\$} \mathsf{Enc}(m_b)$
return c

Let $\mathcal{A}$ *be an adversary. We consider the following experiments:*

Experiment $\mathsf{Exp}^{\text{ind-cpa-1}}_{\mathsf{SE}}(\mathcal{A})$:	Experiment $\mathsf{Exp}^{\text{ind-cpa-0}}_{\mathsf{SE}}(\mathcal{A})$:
$s \xleftarrow{\$} \mathsf{Gen}(1^\lambda)$	$s \xleftarrow{\$} \mathsf{Gen}(1^\lambda)$
$d \xleftarrow{\$} \mathcal{A}^{\mathsf{Enc}(\mathsf{LR}(\cdot,\cdot,1))}$	$d \xleftarrow{\$} \mathcal{A}^{\mathsf{Enc}(\mathsf{LR}(\cdot,\cdot,0))}$
return d	return d

For every PPT $\mathcal{A}$*, there exists a negligible function* ν *such that it holds that:*

$$\left|\Pr[\mathsf{Exp}^{\text{ind-cpa-1}}_{\mathsf{SE}}(\mathcal{A}) = 1] - \Pr[\mathsf{Exp}^{\text{ind-cpa-0}}_{\mathsf{SE}}(\mathcal{A}) = 1]\right| < \nu(\lambda)$$

Definition 8 (Pseudo-Random Function (from book of [29]))**.** *A function* $\mathsf{PRF}\colon \{0,1\}^\lambda \times \{0,1\}^\lambda \to \{0,1\}^{c\lambda}$ *is called a* pseudo-random function *if it satisfies the following properties.* Efficient: *For every* $k \in \{0,1\}^\lambda$*, and every* $m \in \{0,1\}^{c\lambda}$*, there exists a* PPT *algorithm to compute* $\mathsf{PRF}_k(m) = \mathsf{PRF}(k,m)$.

Indistinguishable from Random: *For every* PPT $\mathcal{A}$*, there exists a negligible function* ν*, such that for all auxiliary input* $z \in \{0,1\}^\star$ *it holds that:*

$$\left|\Pr_{k\xleftarrow{\$}\{0,1\}^\lambda}(\mathcal{A}^{\mathsf{PRF}_k}(z) = 1) - \Pr_{f\xleftarrow{\$}F}(\mathcal{A}^{f}(z) = 1)\right| < \nu(\lambda)$$

where $F = \{f\colon \{0,1\}^\lambda \to \{0,1\}^{c\lambda}\}$

Definition 9 (CP-well-formed functionality)**.** *We say that a CP functionality* $\mathcal{F}$ *is a* CP-well-formed functionality *if, when all parties are corrupt* $\mathcal{F}$ *has the following behavior on its adversaries' interfaces:*

1) *Whenever a message* m *is received on the* i*th adversarial interface,* $\mathcal{F}$ *outputs* (i,m) *to the first adversarial interface.*
2) *Whenever a message of the form* $(i,\mathtt{msg})$ *is received on the first adversarial interface,* $\mathcal{F}$ *outputs the message* m *to the* i*th adversarial interface.*

APPENDIX I
UC SECURITY WITH GLOBAL SETUP (GUC)

The core of (GUC) security is the indistinguishability between the real and ideal worlds. In the real world, parties execute a protocol Π and communicate over a channel defined in the model. In the ideal world, parties access a functionality $\mathcal{F}$ which obtains inputs from them and returns to them the output directly. A protocol Π *securely-realizes* a functionality $\mathcal{F}$ if any adversary $\mathcal{A}$ in the real world can be emulated by a simulator Sim in the ideal world. That is, the two worlds cannot be distinguished by a distinguisher, called the environment $\mathcal{Z}$. In addition, the property of *composability* means that a protocol Π remains secure, even after replacing its calls to a subroutine $\mathcal{F}_1$ (a functionality) with calls to a protocol Π_1 that securely realizes $\mathcal{F}_1$. We call Π a $\mathcal{F}_1$-hybrid protocol.

Below, we formally define the GUC framework of Canetti et al. [17]. Let $\mathcal{R}$ and $\bar{\mathcal{G}}$ be functionalities. Let Π be a $\mathcal{R}$-hybrid protocol, $\bar{\mathcal{G}}$ a setup, $\mathcal{A}$ an adversary, and $\mathcal{Z}$ an environment. The output of the environment $\mathcal{Z}$ after an execution of Π in the GUC $\bar{\mathcal{G}}$-hybrid model in presence of $\mathcal{A}$ is denoted as $\text{EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi,\mathcal{A},\mathcal{Z}}$. The output of $\mathcal{Z}$ in the ideal world where the simulator Sim interacts with an ideal functionality $\mathcal{F}$ and the setup $\bar{\mathcal{G}}$ is denoted as $\text{EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\Pi,\mathsf{Sim},\mathcal{Z}}$.

Definition 10 (UC with Global Setup)**.** *Let* $\bar{\mathcal{G}}$ *be a global setup, and* $\mathcal{R}$ *be a resource. For an* n*-party efficient protocol* Π *and functionality* $\mathcal{F}$*, we say that* Π GUC-realizes $\mathcal{F}$ *if* $\forall\mathcal{A}\ \exists\mathsf{Sim}\ \forall\mathcal{Z}$

$$\text{EXEC}^{\bar{\mathcal{G}},\mathcal{R}}_{\Pi,\mathcal{A},\mathcal{Z}} \approx \text{EXEC}^{\bar{\mathcal{G}},\mathcal{F}}_{\mathsf{Sim},\mathcal{Z}}$$

A. The basic signature functionality

In Fig. 9, we provide the basic UC signature functionality proposed in [16] modified to support strong unforgeability.

APPENDIX J
SECURE FUNCTION EVALUATION AND BROADCAST FUNCTIONALITIES

In Fig. 10 and 11 we provide the SFE and broadcast functionalities proposed in [28]. In Fig 12 we provide the formal description of the for unanimous abort functionality

Key Generation.
Upon receiving a value $(\texttt{KEY\_GEN}, \mathsf{sid})$ from some party $S \in \mathcal{P}$, verify that $\mathsf{sid} = (S, \mathsf{sid}')$ for some sid. If not, then ignore the request. Else, hand $(\texttt{KEY\_GEN}, \mathsf{sid})$ to the $\mathcal{A}$. Upon receiving $(\texttt{VERIFICATION\_KEY}, \mathsf{sid}, v)$ from the $\mathcal{A}$, output $(\texttt{VERIFICATION\_KEY}, \mathsf{sid}, v)$ to S, and record the pair (S, v).
Signature. If $I = (\mathsf{SIGN}, \mathsf{sid}, m)$ is received from party S, verify that $\mathsf{sid} = (S, \mathsf{sid}')$ for some sid'. If not, then ignore the request, else send $(\mathsf{SIGN}, \mathsf{sid}, m)$ to $\mathcal{A}$. Upon receiving $I = (\texttt{SIGNATURE}, \mathsf{sid}, m, \sigma)$ from $\mathcal{A}$, verify that no entry $(m, \sigma, v, 0)$ is stored. If it is, then output an error message to S and halt. Else, send $(\texttt{SIGNATURE}, \mathsf{sid}, m, \sigma)$ to S, and store the entry $(m, \sigma, v, 1)$.
Verification
Upon receiving a value $(\texttt{VERIFY}, \mathsf{sid}, m, \sigma, v')$ from some party p_i, hand $(\texttt{VERIFY}, \mathsf{sid}, m, \sigma, v')$ to the adversary. Upon receiving $(\texttt{VERIFIED}, \mathsf{sid}, m, \phi)$ from the adversary do:

1) If $v' = v$ and the entry $(m, \sigma, v, 1)$ is recorded, then set $f = 1$. (This condition guarantees completeness: If the verification key v' is the registered one and σ is a legitimately generated signature for m, then the verification succeeds.)
2) Else, if $v' = v$, the signer is not corrupted, and the entry $(m, \sigma, v, 1)$ is recorded, then set $f = 0$ and record the entry $(m, \sigma, v, 0)$. (This condition guarantees strong unforgeability: If v' is the registered one, the signer is not corrupted, and a signature σ of m has never been generated, then the verification fails.)
3) Else, if there is an entry (m, σ, v', f') stored, then let $f = f'$. (This condition guarantees consistency: All verification requests with identical parameters will result in the same answer.)
4) Else, let $f = \phi$ and record the entry (m, σ, v', ϕ)

Send $(\texttt{VERIFIED}, \mathsf{sid}, m, f)$ to p_i.

Fig. 9. The $\mathcal{F}_{\texttt{SIGN}}$ functionality with strong unforgeability of [16]

$\mathcal{F}_{\mathsf{SFE}}^{f}$ is as follows, given a function $f : (\{0,1\}^* \cup \{\perp\})^n \times R \to (\{0,1\}^*)^n$ and a set of parties $\mathcal{P}$. Initialize the variables $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ to a default value $\perp$.

- Upon receiving (Input, v) from some party $p_i \in \mathcal{P}$, set $x_i := v$ and send a message (Input, i) to the adversary.
- Upon receiving (Output) from some party $p_i \in \mathcal{P}$, do:
  1) If x_j has been set for all $j \in \mathcal{H}$, and $y_1, \ldots, y_n$ have not yet been set, then choose $r \xleftarrow{\$} R$ and set $(y_1, \ldots, y_n) := f(x_1, \ldots, x_n, r)$.
  2) Output y_i to p_i

Fig. 10. The $\mathcal{F}_{\mathsf{SFE}}^{f}$ functionality of [28]

$\mathcal{F}^{\mathsf{UNA}}$ (the description of the functionality extends to 3-party functionality provided in [51]).

$\mathcal{B}$ is as follows, given a set of parties $\mathcal{P}$.

- Upon receiving x_i from party $p_i \in \mathcal{P}$, send (x_i, p_i) to every party in $\mathcal{P}$. (If $\mathcal{B}$ is considered a UC functionality, the output is given in a delayed manner, cf. [15])

Fig. 11. The broadcast functionality $\mathcal{B}$ of [27], [35]

- Upon receiving (Input, v) from some party $p_i \in \mathcal{P}$, set $x_i := v$ and send a message (Input, i) to the adversary. If v is outside the domain of p_i consider $x_i = \texttt{ABORT}$.
- If there exists $i \in \{1, \ldots, n\}$ such that $x_i = \texttt{ABORT}$ the set $(y = \perp)$ else set $y = f(x_1, \ldots, x_n)$ and send y to the adversary.
- Upon receiving $\texttt{ok}$ from the adversary, send y to the honest parties (if they query the functionality to get the output).
- Upon receiving $\texttt{ABORT}$ from the adversary send $\perp$ to the honest parties (if they query the functionality to get the output)

Fig. 12. The $\mathcal{F}^{\mathsf{UNA}}$.