



Financially Backed Covert Security

Sebastian Faust¹, Carmit Hazay², David Kretzler¹, and Benjamin Schlosser¹(✉)

¹ Technical University of Darmstadt, Darmstadt, Germany
{sebastian.f Faust,david.kretzler,benjamin.schlosser}@tu-darmstadt.de
² Bar-Ilan University, Ramat Gan, Israel
carmit.hazay@biu.ac.il

Abstract. The security notion of *covert security* introduced by Aumann and Lindell (TCC’07) allows the adversary to successfully cheat and break security with a fixed probability $1 - \epsilon$, while with probability ϵ , honest parties detect the cheating attempt. Asharov and Orlandi (ASIACRYPT’12) extend covert security to enable parties to create publicly verifiable evidence about misbehavior that can be transferred to any third party. This notion is called *publicly verifiable covert security* (PVC) and has been investigated by multiple works. While these two notions work well in settings with known identities in which parties care about their reputation, they fall short in Internet-like settings where there are only digital identities that can provide some form of anonymity.

In this work, we propose the notion of *financially backed covert security* (FBC), which ensures that the adversary is financially punished if cheating is detected. Next, we present three transformations that turn PVC protocols into FBC protocols. Our protocols provide highly efficient judging, thereby enabling practical judge implementations via smart contracts deployed on a blockchain. In particular, the judge only needs to non-interactively validate a single protocol message while previous PVC protocols required the judge to emulate the whole protocol. Furthermore, by allowing an interactive punishment procedure, we can reduce the amount of validation to a single program instruction, e.g., a gate in a circuit. An interactive punishment, additionally, enables us to create financially backed covert secure protocols without any form of common public transcript, a property that has not been achieved by prior PVC protocols.

Keywords: Covert Security · Multi-Party Computation (MPC) · Public Verifiability · Financial Punishment

1 Introduction

Secure multi-party computation (MPC) protocols allow a set of parties to jointly compute an arbitrary function f on private inputs. These protocols guarantee privacy of inputs and correctness of outputs even if some of the parties are corrupted by an adversary. The two standard adversarial models of MPC are

semi-honest and *malicious* security. While semi-honest adversaries follow the protocol description but try to derive information beyond the output from the interaction, malicious adversaries can behave in an arbitrary way. MPC protocols in the malicious adversary model provide stronger security guarantees at the cost of significantly less efficiency. As a middle ground between good efficiency and high security Aumann and Lindell introduced the notion of *security against covert adversaries* [AL07]. As in the malicious adversary model, corrupted parties may deviate arbitrarily from the protocol specification but the protocol ensures that cheating is detected with a fixed probability, called *deterrence factor* ϵ . The idea of covert security is that adversaries fear to be detected, e.g., due to reputation issues, and thus refrain from cheating.

Although cheating can be detected in covert security, a party of the protocol cannot transfer the knowledge about malicious behavior to other (external) parties. This shortcoming was addressed by Asharov and Orlandi [AO12] with the notion of *covert security with public verifiability* (PVC). Informally, PVC enables honest parties to create a publicly verifiable certificate about the detected malicious behavior. This certificate can subsequently be checked by any other party (often called *judge*), even if this party did not contribute to the protocol execution. The idea behind this notion is to increase the deterrent effect by damaging the reputation of corrupted parties publicly. PVC secure protocols for the two-party case were presented by [AO12, KM15, ZDH19, HKK+19]. Recently, Damgård et al. [DOS20] showed a generic compiler from semi-honest to publicly verifiable covert security for the two-party setting and gave an intuition on how to extend their compiler to the multi-party case. Full specifications of generic compilers from semi-honest to publicly verifiable covert security for multi-party protocols were presented by Faust et al. [FHKS21] and Scholl et al. [SSS21].

Although PVC seems to solve the shortcoming of covert security at first glance, in many settings PVC is not sufficient; especially, if only a digital identity of the parties is known, e.g., in the Internet. In such a setting, a real party can create a new identity without suffering from a damaged reputation in the sequel. Hence, malicious behavior needs to be punished in a different way. A promising approach is to use existing cryptocurrencies to directly link cheating detection to financial punishment without involving trusted third parties; in particular, cryptocurrencies that support so-called *smart contracts*, i.e., programs that enable the transfer of assets based on predefined rules. Similar to PVC, where an external judge verifies cheating by checking a certificate of misbehavior, we envision a smart contract that decides whether a party behaved maliciously or not. In this setting, the task of judging is executed over a distributed blockchain network keeping it incorruptible and verifiable at the same time. Since every instruction executed by a smart contract costs fees, it is highly important to keep the amount of computation performed by a contract small. This aspect is not solely important for execution of smart contracts but in all settings where an external judge charges by the size of the task it gets. Due to this constraint, we cannot straightforwardly adapt PVC protocols to work in this setting, since detection of malicious behavior in existing PVC protocols is performed in a naive way that requires the judge to recompute a whole protocol execution.

Related Work. While combining MPC with blockchain technologies is an active research area (e.g., [KB14, BK14, ADMM14]) none of these works deal with realizing the judging process of PVC protocols over a blockchain. The only work connecting covert security with financial punishment thus far is by Zhu et al. [ZDH19], which we describe in a bit more detail below. They combine a two-party garbling protocol with an efficient judge that can be realized via a smart contract. Their construction leverages strong security primitives, like a malicious secure oblivious transfer for the transmission of input wires, to ensure that cheating can only occur during the transmission of the garbled circuit and not in any other part of the two-party protocol. By using a binary search over the transmitted circuit, the parties narrow down the computation step under dispute to a single circuit gate. This process requires $O(\log(|C|))$ interactions, where $|C|$ denotes the circuit size, and enables the judge to resolve the dispute by recomputing only a single circuit gate.

While the approach of Zhu et al. [ZDH19] provides an elegant way to reduce the computational complexity of the judge in case cheating is restricted to a single message, it falls short if multiple messages or even a whole protocol execution is under dispute. As a consequence, their construction is limited in scalability and generality, since it is only applicable to two-party garbling protocols, i.e., neither other semi-honest two-party protocols nor more parties are supported.

Generalizing the ideas of [ZDH19] to work for other protocol types and the multi-party case requires us to address several challenges. First, in [ZDH19] the transmitted garbled circuit under dispute is the result of the completely non-interactive garbling process. In contrast, many semi-honest MPC protocols (e.g., [GMW87, BMR90]) consist of several rounds of interactions that need to be all considered during the verification. Interactivity poses the challenge that multiple messages may be under dispute and the computation of messages performed by parties may depend on data received in previous rounds. Hence, verifications of messages need to consider local computations and internal states of the parties that depend on all previous communication rounds. This task is far more complex than verifying a single public message. Second, supporting more than two parties poses the challenge of resolving a dispute about a protocol execution during which parties might not know the messages sent between a subset of other parties. Third, the transmitted garbled circuit in [ZDH19] is independent of the parties private inputs. Considering protocols where parties provide secret inputs or messages that depend on these inputs, requires a privacy-preserving verification mechanism to protect parties' sensitive data.

1.1 Contribution

Our first contribution is to introduce a new security notion called *financially backed covert security* (FBC). This notion combines a covertly secure protocol with a mechanism to financially punish a corrupted party if cheating was

detected. We formalize financial security by adding two properties to covert security, i.e., *financial accountability* and *financial defamation freeness*. Our notion is similar to the one of PVC; in fact, PVC adds reputational punishment to covert security via *accountability* and *defamation freeness*. In order to lift these properties to the financial context, FBC requires deposits from all parties and allows for an interactive judge. We present two security games to formalize our introduced properties. While the properties are close to accountability and defamation freeness of PVC, our work for the first time explicitly presents formal security games for these security properties, thereby enabling us to rigorously reason about financial properties in PVC protocols. We briefly compare our new notion to the security definition of Zhu et al. [ZDH19], which is called *financially secure computation*. Zhu et al. follow the approach of simulation-based security by presenting an ideal functionality for two parties that extends the ideal functionality of covert security. In contrast, we present a game-based security definition that is not restricted to the two-party case. While simulation-based definitions have the advantage of providing security under composition, proving a protocol secure under their notion requires to create a full simulation proof which is an expensive task. Instead, our game-based notion allows to re-use simulation proofs of all existing covert and PVC protocols, including future constructions, and to focus on proving financial accountability and financial defamation freeness in a standalone way.

We present transformations from different classes of PVC protocols to FBC protocols. While we could base our transformations on covert protocols, FBC protocols require a property called *prevention of detection dependent abort*, which is not always guaranteed by a covert protocol. The property ensures that a corrupted party cannot abort after learning that her cheating will be detected without leaving publicly verifiable evidence. PVC protocols always satisfy prevention of detection dependent abort. So, by basing our transformation on PVC protocols, we inherit this property.

While the mechanism utilized by [ZDH19] to validate misbehavior is highly efficient, it has only been used for non-interactive algorithms so far, i.e., to validate correctness of the garbling process. We face the challenge of extending this mechanism over an interactive protocol execution while still allowing for efficient dispute resolution such that the judge can be realized via a smart contract. In order to tackle these challenges, we present a novel technique that enables efficient validation of arbitrary complex and interactive protocols given the randomness and inputs of all parties. What's more, we can allow for private inputs if a public transcript of all protocol messages is available. We utilize only standard cryptographic primitives, in particular, commitments and signatures.

We differentiate existing PVC protocols according to whether the parties provide private inputs or not. The former protocols are called *input-dependent* and the latter ones *input-independent*. Input-independent protocols are typically used to generate correlated randomness. Further, all existing PVC protocols incorporate some form of common public transcript. Input-dependent protocols require a common public transcript of messages. In contrast, for input-independent pro-

protocols, it is enough to agree on the hashes of all sent messages. While it is not clear, if it is possible to construct PVC protocols without any form of public transcript, we construct FBC protocols providing this property. We achieve this by exploiting the interactivity of the judge, which is non-interactive in PVC. Based on the above observations, we define the following three classes of FBC protocols, for which we present transformations from PVC protocols.

- Class 1:** The first class contains *input-independent* protocols during which parties learn hashes of all protocol messages such that they agree on a common *transcript of message hashes*.
- Class 2:** The second class contains *input-dependent* protocols with a public *transcript of messages*. In contrast to class 1, parties may provide secret inputs and share a common view on all messages instead of a common view on hashes only.
- Class 3:** The third class contains input-independent protocols where parties do not learn any information about messages exchanged between a subset of other parties (cf. class 1). As there are no PVC protocol fitting into this class, we first convert PVC protocols matching the requirements of class 1 into protocols without public transcripts and second leverage an interactive punishment procedure to transform the resulting protocols into FBC protocols without public transcripts. Our FBC protocols benefit from this property since parties have to send all messages only to the receiver and not to all other parties. This effectively reduces the concrete communication complexity by a factor depending on the number of parties. In the optimistic case, if there is no cheating, we get this benefit without any overhead in the round complexity.

For each of our constructions, we provide a formal specification and a rigorous security analysis; the ones of the second class can be found in the full version of this paper. This is in contrast to the work of [ZDH19] which lacks a formal security analysis for financially secure computation. We stress that all existing PVC multi-party protocols can be categorized into class 1 and 2. Additionally, by combining any of the transformations from [DOS20, FHKS21, SSS21], which compile semi-honest protocols into PVC protocols, our constructions can be used to transform these protocol into FBC protocols.

The resulting FBC protocols for class 1 and 2 allow parties to non-interactively send evidence about malicious behavior to the judge. As the judge entity in these two classes is non-interactive, techniques from our transformations are of independent interest to make PVC protocols more efficient. Since, in contrast to class 1 and 2, there is no public transcript present in protocols of class 3, we design an interactive process involving the judge entity to generate evidence about malicious behavior. For all protocols, once the evidence is interactively or non-interactively created, the judge can efficiently resolve the dispute by recomputing only a single protocol message regardless of the overall computation size. We can further reduce the amount of validation to a single program instruction, e.g., a gate in a circuit, by prepending an interactive search procedure. This extension is presented in the full version of this paper.

Finally, we provide a smart contract implementation of the judging party in Ethereum and evaluate its gas costs (cf. Sect. 8). The evaluation shows the practicability, e.g., in the three party setting, with optimistic execution costs of 533 k gas. Moreover, we show that the dispute resolution of our solution is highly scalable in regard to the number of parties, the number of protocol rounds and the protocol complexity.

1.2 Technical Overview

In this section, we outline the main techniques used in our work and present the high-level ideas incorporated into our constructions. We start with an overview of the new notion of *financially backed covert security*. Then, we present a first attempt of a construction over a blockchain and outline the major challenges. Next, we describe the main techniques used in our constructions for PVC protocols of classes 1 and 2 and finally elaborate on the bisection procedure required for the more challenging class 3.

Financially Backed Covert Security. We recall that, a publicly verifiable covertly secure (PVC) protocol $(\pi_{\text{cov}}, \text{Blame}, \text{Judge})$ consists of a covertly secure protocol π_{cov} , a blaming algorithm **Blame** and a judging algorithm **Judge**. The blaming algorithm produces a certificate *cert* in case cheating was detected and the judging algorithm, upon receiving a valid certificate, outputs the identity of the corrupted party. The algorithm **Judge** of a PVC protocol is explicitly defined as non-interactive. Therefore, *cert* can be transferred at any point in time to any third party that executes **Judge** and can be convinced about malicious behavior if the algorithm outputs the identity of a corrupted party.

In contrast to PVC, *financially backed covert security* (FBC) works in a model where parties own assets which can be transferred to other parties. This is modelled via a ledger entity \mathcal{L} . Moreover, the model contains a trusted judging party \mathcal{J} which receives deposits before the start of the protocol and adjudicates in case of detected cheating. We emphasize that the entity \mathcal{J} , which is a single trusted entity interacting with all parties, is not the same as the algorithm **Judge** of a PVC protocol, which can be executed non-interactively by any party. An FBC protocol $(\pi'_{\text{cov}}, \text{Blame}', \text{Punish})$ consists of a covertly secure protocol π'_{cov} , a blaming algorithm **Blame'** and an interactive punishment protocol **Punish**. Similar to PVC, the blaming algorithm **Blame'** produces a certificate *cert'* that is used as an input to the interactive punishment protocol. **Punish** is executed between the parties and the judge \mathcal{J} . If all parties behave honestly during the execution of π'_{cov} , \mathcal{J} sends the deposited coins back to all parties after the execution of **Punish**. In case cheating is detected during π'_{cov} , the judge \mathcal{J} burns the coins of the cheating party.

First Attempt of an Instantiation Over a Blockchain. Blockchain technologies provide a convenient way of handling monetary assets. In particular, in combination with the execution of smart contracts, e.g., offered by Ethereum [W+14], we envision to realize the judging party \mathcal{J} as a smart contract. A first attempt

of designing the punishment protocol is to implement \mathcal{J} in a way, that the judge just gets the certificate generated by the PVC protocol’s blame algorithm and executes the PVC protocol’s **Judge**-algorithm. However, the **Judge**-algorithm of all existing PVC protocols recomputes a whole protocol instance and compares the output with a common transcript on which all parties agree beforehand. As computation of a smart contract costs money in form of transaction fees, recomputing a whole protocol is prohibitively expensive. Therefore, instead of recomputing the whole protocol, we aim for a punishment protocol that facilitates a judging party \mathcal{J} which needs to recompute just a single protocol step or even a single program instruction, e.g., a gate in a circuit. The resulting judge becomes efficient in a way that it can be practically realized via a smart contract.

FBC Protocols with Efficient Judging from PVC Protocols. In this work, we present three transformations from PVC protocols to FBC protocols. Our transformations start with PVC protocols providing different properties which we use to categorize these protocols into three classes. We model the protocol execution in a way such that every party’s behavior is deterministically defined by her input, her randomness and incoming messages. More precisely, we define the initial state of a party as her input and some randomness and compute the next state according to the state of the previous round and the incoming messages of the current round. Our first two transformations build on PVC protocols where the parties share a public transcript of the exchanged messages resp. message hashes. Additionally, parties send signed commitments on their intermediate states to all parties. The opening procedure ensures that correctly created commitments can be opened – falsely created commitments open to an invalid state that is interpreted as an invalid message. By sending the internal state of some party P_m for a single round together with the messages received by P_m in the same round to the judging party, the latter can efficiently verify malicious behavior by recomputing just a single protocol step. The resulting punishment protocol is efficient and can be executed without contribution of the cheating party.

Interactive Punishment Protocol to Support Private Transcripts. Our third transformation compiles input-independent PVC protocols with a public transcript into protocols where no public transcript is known to the parties. The lack of a public transcript makes the punishment protocol more complicated. Intuitively, since an honest party has no signed information about the message transcript, she cannot provide verifiable data about the incoming message used to calculate a protocol step. Therefore, we use the technique of an interactive bisection protocol which was first used in the context of verifiable computing by Canetti et al. [CRR11] and subsequently by many further constructions [KGC+18, TR19, ZDH19, EFS20]. While the bisection technique is very efficient to narrow down disagreement, it was only used for non-interactive algorithms so far. Hence, we extend this technique to support also interactive protocols. In particular, in our work, we use a bisection protocol to allow two parties to efficiently agree on a common message history. To this end, both parties, the

accusing and the accused one, create a Merkle tree of their emulated message history up to the disputed message and submit the corresponding root. If they agree on the message history, the accusation can be validated by reference to this history. If they disagree, they perform a bisection search over the proposed history that determines the first message in the message history, they disagree on, while automatically ensuring that they agree on all previous messages. Hence, the judge can verify the message that the parties disagree on based on the previous messages they agree on. At the end of both interactions, the judge can efficiently resolve the dispute by recomputing just a single step.

2 Preliminaries

We start by introducing notation and cryptographic primitives used in our construction. Moreover, we provide the definition of covert security and publicly verifiable covert security in the full version of this paper.

We denote the computational security parameter by κ . Let n be some integer, then $[n] = \{1, \dots, n\}$. Let $i \in [n]$, then we use the notation $j \neq i$ for $j \in [n] \setminus \{i\}$. A function $\text{negl}(n) : \mathbb{N} \rightarrow \mathbb{R}$ is *negligible* in n if for every positive integer c there exists an integer n_0 such that $\forall n > n_0$ it holds that $\text{negl}(n) < \frac{1}{n^c}$. We use the notation $\text{negl}(n)$ to denote a negligible function.

We define $\text{REAL}_{\pi, \mathcal{A}(z), \mathcal{I}}(\bar{x}, 1^\kappa)$ to be the output of the execution of an n -party protocol π executed between parties $\{P_i\}_{i \in [n]}$ on input $\bar{x} = \{x_i\}_{i \in [n]}$ and security parameter κ , where \mathcal{A} on auxiliary input z corrupts parties $\mathcal{I} \subset \{P_i\}_{i \in [n]}$. We further specify $\text{OUTPUT}_j(\text{REAL}_{\pi, \mathcal{A}(z), \mathcal{I}}(\bar{x}, 1^\kappa))$ to be the output of party P_j for $j \in [n]$.

Our protocol utilizes a signature scheme (**Generate**, **Sign**, **Verify**) that is *existentially unforgeable under chosen-message attacks*. We assume that each party executes the **Generate**-algorithm to obtain a key pair (pk, sk) before the protocol execution. Further, we assume that all public keys are published and known to all parties while the secret keys are kept private. To simplify the protocol description we denote signed messages with $\langle x \rangle_i$ instead of $(x, \sigma := \text{Sign}_{\text{sk}_i}(x))$. The verification is therefore written as $\text{Verify}(\langle x \rangle_i)$ instead of $\text{Verify}_{\text{pk}_i}(x, \sigma)$. Further, we make use of a hash function $H(\cdot) : \{0, 1\}^* \rightarrow \{0, 1\}^\kappa$ that is collision resistant.

We assume a synchronous communication model, where communication happens in rounds and all parties are aware of the current round. Messages that are sent in some round k arrive at the receiver in round $k + 1$. Since we consider a rushing adversary, the adversary learns the messages sent by honest parties in round k in the same round and hence can adapt her own messages accordingly. We denote a message sent from party P_i to party P_j in round k of some protocol instance denoted with ℓ as $\text{msg}_{(\ell, k)}^{(i, j)}$. The hash of this message is denoted with $\text{hash}_{(\ell, k)}^{(i, j)} := H(\text{msg}_{(\ell, k)}^{(i, j)})$.

A *Merkle tree* over an ordered set of elements $\{x_i\}_{i \in [N]}$ is a labeled binary hash tree, where the i -th leaf is labeled by x_i . We assume N to be an integer power of two. In case the number of elements is not a power of two, the set can

be padded until N is a power of two. For construction of Merkle trees, we make use of the collision-resistant hash function $H(\cdot) : \{0, 1\}^* \rightarrow \{0, 1\}^\kappa$.

Formally, we define a Merkle tree as a tuple of algorithms (MTree, MRoot, MProof, MVerify). Algorithm MTree takes as input a computational security parameter κ as well as a set of elements $\{x_i\}_{i \in [N]}$ and creates a Merkle tree \mathbf{mTree} . To ease the notation, we will omit the security parameter and implicitly assume it to be provided. Algorithm MRoot takes as input a Merkle tree \mathbf{mTree} and returns the root element root of tree \mathbf{mTree} . Algorithm MProof takes as input a leaf x_j and Merkle tree \mathbf{mTree} and creates a Merkle proof σ showing that x_j is the j -th leaf in \mathbf{mTree} . Algorithm MVerify takes as input a proof σ , an index i , a root root and a leaf x^* and returns true iff x^* is the i -th leaf of a Merkle tree with root root .

A Merkle Tree satisfies the following two requirements. First, for each Merkle tree \mathbf{mTree} created over an arbitrary set of elements $\{x_i\}_{i \in [N]}$, it holds that for each $j \in [N]$ $\text{MVerify}(\text{MProof}(x_j, \mathbf{mTree}), j, \text{MRoot}(\mathbf{mTree}), x_j) = \text{true}$. We call this property *correctness*. Second, for each Merkle tree \mathbf{mTree} with root $\text{root} := \text{MRoot}(\mathbf{mTree})$ created over an arbitrary set of elements $\{x_i\}_{i \in [N]}$ with security parameter κ it holds that for each polynomial time algorithm adversary \mathcal{A} outputting an index j^* , leaf $x^* \neq x_{j^*}$ and proof σ^* the probability that $\text{MVerify}(\sigma^*, j^*, \text{MRoot}(\mathbf{mTree}), x^*) = \text{true}$ is $\text{negl}(\kappa)$. We call this property *binding*.

3 Financially Backed Covert Security

In the following, we specify the new notion of *financially backed covert security*. This notion extends covert security by a mechanism of financial punishment. More precisely, once an honest party detects cheating of the adversary during the execution of the covertly secure protocol, there is some corrupted party that is financial punished afterwards. The financial punishment is realized by an interactive protocol Punish that is executed directly after the covertly secure protocol. In order to deal with monetary assets, financially backed covertly secure protocols depend on a public ledger \mathcal{L} and a trusted judge \mathcal{J} . The former can be realized by distributed ledger technologies, such as blockchains, and the latter by a smart contract executed on the said ledger. In the following, we describe the role of the ledger and the judging party, formally define financially backed covert security and outline techniques to prove financially backed covert security.

3.1 The Ledger and Judge

An inherent property of our model is the handling of assets and asset transfers based on predefined conditions. Nowadays, distributed ledger technologies like blockchains provide convenient means to realize this functionality. We model the handling of assets resp. coins via a ledger entity denoted by \mathcal{L} . The entity stores a balance of coins for each party and transfers coins between parties upon request. More precisely, \mathcal{L} stores a balance $b_i^{(t)}$ for each party P_i at time t . For

the security definition presented in Sect. 3.2, we are in particular interested in the balances before the execution of the protocol π , i.e., $b_i^{(\text{pre})}$, and after the execution of the protocol **Punish**, i.e., $b_i^{(\text{post})}$. The balances are public such that every party can query the amount of coins for any party at the current time. In order to send coins to another party, a party interacts with \mathcal{L} to trigger the transfer.

While we consider the ledger as a pure storage of balances, we realize the conditional transfer of coins based on some predefined rules specified by the protocol **Punish** via a judge \mathcal{J} . In particular, \mathcal{J} constitutes a trusted third party that interacts with the parties of the covertly secure protocol. More precisely, we require that each party sends some fixed amount of coins as deposit to \mathcal{J} before the covertly secure protocol starts. During the covertly secure protocol execution, the judge keeps the deposited coins but does not need to be part of any interaction. After the execution of the covertly secure protocol, the judge plays an important role in the punishment protocol **Punish**. In case any party detects cheating during the execution of the covertly secure protocol, \mathcal{J} acts as an adjudicator. If there is verifiable evidence about malicious behavior of some party, the judge financially punishes the corrupted party by withholding her deposit. Eventually, \mathcal{J} will reimburse all parties with their deposits except those parties that have been proven to be malicious. The rules according to which parties are considered malicious and hence according to which the coins are reimbursed or withheld need to be specified by the protocol **Punish**.

Finally, we emphasize that both entities the ledger \mathcal{L} and the judge \mathcal{J} are considered trusted. This means, the correct functionality of these entities cannot be distorted by the adversary.

3.2 Formal Definition

We work in a model in which a ledger \mathcal{L} and a judge \mathcal{J} as explained above exist. Let π' be an n -party protocol that is covertly secure with deterrence factor ϵ . Let the number of corrupted parties that is tolerated by π' be $m < n$ and the set of corrupted parties be denoted by \mathcal{I} . We define π as an extension of π' , in which all involved parties transfer a fixed amount of coins, d , to \mathcal{J} before executing π' . Additionally, after the execution of π' , all parties execute algorithm **Blame** which on input the view of the honest party outputs a certificate and broadcasts the generated certificate – still as part of π . The certificate is used for both proving malicious behavior, if detected, and defending against being accused for malicious behavior.

After the execution of π , all parties participate in the protocol **Punish**. In case honest parties detected misbehavior, they prove said misbehavior to \mathcal{J} such that \mathcal{J} can punish the malicious party. In case a malicious party blames an honest one, the honest parties participate to prove their correct behavior. Either way, even if there is no blame at all, all honest parties wait to receive their deposits back, which are reimbursed by \mathcal{J} at the end of the punishment protocol **Punish**.

Definition 1 (Financially backed covert security). We call a triple $(\pi, \text{Blame}, \text{Punish})$ an n -party financially backed covertly secure protocol with deterrence factor ϵ computing some function f in the \mathcal{L} and \mathcal{J} model, if the following security properties are satisfied:

1. **Simulatability with ϵ -deterrent:** The protocol π (as described above) is secure against a covert adversary according to the strong explicit cheat formulation with ϵ -deterrent and non-halting detection accurate.
2. **Financial Accountability:** For every PPT adversary \mathcal{A} corrupting parties P_i for $i \in \mathcal{I} \subset [n]$, there exists a negligible function $\mu(\cdot)$ such that for all $(\bar{x}, z) \in (\{0, 1\})^{n+1}$ the following holds:
If for any honest party $P_h \in [n] \setminus \mathcal{I}$ it holds that $\text{OUTPUT}_h(\text{REAL}_{\pi, \mathcal{A}(z), \mathcal{I}}(\bar{x}, 1^\kappa)) = \text{corrupted}_*^1$, then $\exists m \in \mathcal{I}$ such that:

$$\Pr[b_m^{(\text{post})} = b_m^{(\text{pre})} - d] > 1 - \mu(\kappa),$$

where d denotes the amount of deposited coins per party.

3. **Financial Defamation Freeness:** For every PPT adversary \mathcal{A} corrupting parties P_i for $i \in \mathcal{I} \subset [n]$, there exists a negligible function $\mu(\cdot)$ such that for all $(\bar{x}, z) \in (\{0, 1\})^{n+1}$ and all $j \in [n] \setminus \mathcal{I}$ the following holds:

$$\Pr[b_j^{(\text{post})} < b_j^{(\text{pre})}] < \mu(\kappa).$$

Remark 1. For simplicity, we assume that the adversary does not transfer coins after sending the deposit to \mathcal{J} . This assumption can be circumvented by restating financial accountability such that the sum of the balances of all corrupted parties (not just the ones involved in the protocol) is reduced by d .

3.3 Proving Security of Financially Backed Covert Security

Our notion of financially backed covert security (FBC) consists of three properties. The simulatability property requires the protocol π , which augments the covertly secure protocol π' , to be covertly secure as well. This does not automatically follow from the security of π' , in particular since π includes the broadcast of certificates in case of detected cheating. Showing simulatability of π guarantees that the adversary does not learn sensitive information from the certificates. Showing that a protocol π satisfies the simulatability property is proven via a simulation proof. In contrast, we follow a game-based approach to formally prove financial accountability and financial defamation freeness. To this end, we introduce two novel security games, Exp^{FA} and Exp^{DF} , in the following. Although these two properties are similar to the accountability and defamation freeness properties of PVC, we are the first to introduce formal security games for any of these properties. While we focus on financial accountability and financial

¹ We use the notation corrupted_* to denote that the output of P_h is corrupted_i for some $i \in \mathcal{I}$. We stress that i does not need to be equal to m of the financial accountability property.

defamation freeness, we note that our approach and our security games can be adapted to suit for the security properties of PVC as well.

Both security games are played between a challenger \mathcal{C} and an adversary \mathcal{A} . We define the games in a way that allows us to abstract away most of the details of π . In particular, we parameterize the games by two inputs, one for the challenger and one for the adversary. The challenger's input contains the certificates $\{\text{cert}_i\}_{i \in [n] \setminus \mathcal{I}}$ of all honest parties generated by the Blame-algorithm after the execution of π while the adversary's input consists of all malicious parties' views $\{\text{view}_i\}_{i \in \mathcal{I}}$. By introducing the certificates as inputs to the game, we can prove financial accountability and financial defamation freeness independent from proving simulatability of protocol π .

Throughout the execution of the security games, the adversary executes one instance of the punishment protocol **Punish** with the challenger that takes over the roles of all honest and trusted parties, i.e., the honest protocol parties P_h for $h \notin \mathcal{I}$, the judge \mathcal{J} , and the ledger \mathcal{L} . To avoid an overly complex challenger description, we define those parties as separated entities that can be addressed by the adversary separately and are all executed by the challenger: $\{P_h\}_{h \in [n] \setminus \mathcal{I}}$, \mathcal{J} , and \mathcal{L} . In case any entity is supposed to act pro-actively and does not only wait to react to malicious behavior, the entity is invoked by the challenger. Communication between said entities is simulated by the challenger. The adversary acts on behalf of the corrupted parties.

Financial Accountability Game. Intuitively, financial accountability states that whenever any honest party detects cheating, there is some corrupted party that loses her deposit. Therefore, we require that the output of all honest parties was **corrupted** $_m$ for $m \in \mathcal{I}$ in the execution of π . If this holds, the security game executes **Punish** as specified by the FBC protocol. Before the execution of **Punish**, the challenger asks the ledger for the balances of all parties and stores them as $\{b_i^{(\text{prePunish})}\}_{i \in [n]}$. Note that **prePunish** denotes the time before **Punish** but after the whole protocol already started. This means, relating to Definition 1, the security deposits are already transferred to \mathcal{J} , i.e., $b_i^{\text{prePunish}} = b_i^{\text{pre}} - d$. After the execution, the challenger \mathcal{C} again reads the balances of all parties storing them as $\{b_i^{(\text{post})}\}_{i \in [n]}$. If $b_m^{(\text{post})} = b_m^{(\text{prePunish})} + d$ for all $m \in \mathcal{I}$, i.e., all corrupted parties get their deposits back, the adversary wins and \mathcal{C} outputs 1, otherwise \mathcal{C} outputs 0. A protocol satisfies the financial accountability property as stated in Definition 1 if for each adversary \mathcal{A} running in time polynomial in κ the probability that \mathcal{A} wins game Exp^{FA} is at most negligible, i.e., if $\Pr[\text{Exp}^{\text{FA}}(\mathcal{A}, \kappa) = 1] \leq \text{negl}(\kappa)$.

Financial Defamation Freeness Game. Intuitively, financial defamation freeness states that an honest party can never lose her deposit as a result of executing the **Punish** protocol. The security game is executed in the same way as the financial accountability game. It only differs in the winning conditions for the adversary. After the execution \mathcal{C} checks the balances of the honest parties. If $b_h^{(\text{post})} < b_h^{(\text{prePunish})} + d$ for at least one $h \in [n] \setminus \mathcal{I}$, the adversary wins and the challenger outputs 1, otherwise \mathcal{C} outputs 0. A protocol satisfies the financial

defamation freeness property as stated in Definition 1 if for each adversary \mathcal{A} running in time polynomial in κ the probability that \mathcal{A} wins game Exp^{DF} is at most negligible, i.e. if $\Pr[\text{Exp}^{\text{DF}}(\mathcal{A}, \kappa) = 1] \leq \text{negl}(\kappa)$.

4 Features of PVC Protocols

We present transformations from different classes of *publicly verifiable covertly secure* multi-party protocols (PVC) to *financially backed covertly secure* protocols (FBC). As our transformations make use of concrete features of the PVC protocol (e.g., the exchanged messages), we cannot use the PVC protocol in a block-box way. Instead, we model the PVC protocol in an abstract way, stating features that are required by our constructions. In the remainder of this section, we present these features in detail and describe how we model them. We note that all existing PVC multi-party protocols [DOS20, FHKS21, SSS21] provide the features specified in this section.

4.1 Cut-and-Choose

Although not required per definition of PVC, a fundamental technique used by all existing PVC protocols is the *cut-and-choose* approach that leverages a semi-honest protocol by executing t instances of the semi-honest protocol in parallel. Afterwards, the views (i.e., input and randomness) of the parties is revealed in s instances. This enables parties to detect misbehavior with probability $\epsilon = \frac{s}{t}$. PVC protocols can be split into protocols where parties provide private inputs and those where parties do not have secret data. While cut-and-choose for input-independent protocols, i.e., those where parties do not have private inputs, work as explained on a high level before, the approach must be utilized in such a way that input privacy is guaranteed for input-dependent protocols. However, for both classes of protocols, a cheat detection probability of $\epsilon = \frac{s}{t}$ can be achieved. We elaborate more on the two variants and provide details about them in the full version of this paper.

4.2 Verification of Protocol Executions

An important feature of PVC protocols based on cut-and-choose is to enable parties to verify the execution of the opened protocol instances. This requires parties to emulate the protocol messages and compare them with the messages exchanged during the real execution. In order to emulate honest behavior, we need the protocol to be derandomized.

Derandomization of the Protocol Execution. In general, the behavior of each party during some protocol execution depends on the party’s private input, its random tape and all incoming messages. In order to enable parties to check the behavior of other parties in retrospect, the actions of all parties need to be made deterministic. To this end, we require the feature of a PVC protocol that all

random choices of a party P_i in a protocol instance are derived from some random seed seed_i using a pseudorandom generator (PRG). The seed seed_i is fixed before the beginning of the execution. It follows that the generated outgoing messages are computed deterministically given the seed seed_i , the secret input and all incoming messages.

State Evolution. Corresponding to our communication model (cf. Sect. 2), the internal states of the parties in a semi-honest protocol instance evolve in rounds. For each party P_i , for $i \in [n]$, and each round $k > 0$ the protocol defines a state transition computeRound_k^i that on input the previous internal state $\text{state}_{(k-1)}^{(i)}$ and the set of incoming messages $\{\text{msg}_{(k-1)}^{(j,i)}\}_{j \neq i}$ computes the new internal state $\text{state}_{(k)}^{(i)}$ and the set of outgoing messages $\{\text{msg}_{(k)}^{(i,j)}\}_{j \neq i}$. Based on the derandomization feature, the state transition is deterministic, i.e., all random choices are derived from a random seed included in the internal state of a party. Each party starts with an initial internal state that equals its random seed seed_i and its secret input x_i . In case no secret input is present (i.e., in the input-independent setting) or no message is sent, the value is considered to be a dummy symbol (\perp). We denote the set of all messages sent during a protocol instance by *protocol transcript*. Summarizing, we formally define

$$\begin{aligned} \text{state}_{(0)}^{(i)} &\leftarrow (\text{seed}_i, x_i) \\ \{\text{msg}_{(0)}^{(j,i)}\}_{j \in [n] \setminus \{i\}} &\leftarrow \{\perp\}_{j \in [n] \setminus \{i\}} \\ (\text{state}_{(k)}^{(i)}, \{\text{msg}_{(k)}^{(i,j)}\}_{j \in [n] \setminus \{i\}}) &\leftarrow \text{computeRound}_k^i(\text{state}_{(k-1)}^{(i)}, \{\text{msg}_{(k-1)}^{(j,i)}\}_{j \in [n] \setminus \{i\}}). \end{aligned}$$

Protocol Emulation. In order to check for malicious behavior, parties locally emulate the protocol execution of the opened instances and compare the set of computed messages with the received ones. In case some involved parties are not checked (e.g., in the input-dependent setting), the emulation gets their messages as input and assumes them to be correct. In this case, in order to ensure that each party can run the emulation, it is necessary that each party has access to all messages sent in the opened instance (cf. Sect. 4.4).

To formalize the protocol emulation, we define for each n -party protocol π with R rounds two emulation algorithms. The first algorithm $\text{emulate}_\pi^{\text{full}}$ emulates all parties while the second algorithm $\text{emulate}_\pi^{\text{part}}$ emulates only a partial subset of the parties and considers the messages of all other parties as correct. We formally define them as

$$\begin{aligned} (\{\text{msg}_{(k)}^{(i,j)}\}_{k,i,j \neq i}, \{\text{state}_{(k)}^{(i)}\}_{k,i}) &\leftarrow \text{emulate}_\pi^{\text{full}}(\{\text{state}_{(0)}^{(i)}\}_i) \quad \text{and} \\ (\{\text{msg}_{(k)}^{(i,j)}\}_{k,i,j \neq i}, \{\text{state}_{(k)}^{(i)}\}_{k,\hat{i}}) &\leftarrow \text{emulate}_\pi^{\text{part}}(O, \{\text{state}_{(0)}^{(i)}\}_{\hat{i}}, \{\text{msg}_{(k)}^{(i^*,j)}\}_{k,i^*,j \neq i^*}) \end{aligned}$$

where $k \in [R]$, $i, j \in [n]$, $\hat{i} \in O$ and $i^* \in [n] \setminus O$. O denotes the set of opened parties.

4.3 Deriving the Initial States

As a third feature, we require a mechanism for the parties of a PVC protocol to learn the initial states of all opened parties in order to perform the protocol emulation (cf. Sect. 4.2). Since PVC prevents detection dependent abort, parties learn the initial state even if the adversary aborts after having learned the cut-and-choose selection. Existing multi-party PVC protocols provide this feature by either making use of oblivious transfer or time-lock puzzles as in [DOS20] resp. [FHKS21, SSS21]. We elaborate on these protocols in the full version of this paper.

To model this behavior formally, we define the abstract tuples $\text{initData}^{\text{core}}$ and $\text{initData}^{\text{aux}}$ as well as the algorithm derivelnit . $\text{initData}_{(i)}^{\text{core}}$ represents data each party holds that should be signed by P_i and can be used to derive the initial state of party P_i in a single protocol instance (e.g., a signed time-lock puzzle). $\text{initData}_{(i)}^{\text{aux}}$ represents the additional data all parties receive during the PVC protocol that can be used to interpret $\text{initData}_{(i)}^{\text{core}}$ (e.g., the verifiable solution of the time-lock puzzle). Finally, derivelnit is an algorithm that on input $\text{initData}_{(i)}^{\text{core}}$ and $\text{initData}_{(i)}^{\text{aux}}$ derives the initial state of party P_i (e.g., verifying the solution of the puzzle). Instead of outputting an initial state, the algorithm derivelnit can also output bad or \perp . The former states that party P_i misbehaved during the PVC protocol by providing inconsistent data. The symbol \perp states that the input to derivelnit has been invalid which can only occur if $\text{initData}_{(i)}^{\text{core}}$ or $\text{initData}_{(i)}^{\text{aux}}$ have been manipulated.

Similar to commitment schemes, our abstraction satisfies a *binding* and *hiding* requirement, i.e., it is computationally *binding* and computationally *hiding*. The binding property requires that the probability of any polynomial time adversary finding a tuple (x, y_1, y_2) such that $\text{derivelnit}(x, y_1) \neq \perp$, $\text{derivelnit}(x, y_2) \neq \perp$, and $\text{derivelnit}(x, y_1) \neq \text{derivelnit}(x, y_2)$ is negligible. The hiding property requires that the probability of a polynomial time adversary finding for a given $\text{initData}^{\text{core}}$ a $\text{initData}^{\text{aux}}$ such that $\text{derivelnit}(\text{initData}^{\text{core}}, \text{initData}^{\text{aux}}) \neq \perp$ is negligible.

4.4 Public Transcript

A final feature required by PVC protocols of class 1 and 2 is the availability of a common public transcript. We define three levels of transcript availability. First, a *common public transcript of messages* ensures that all parties hold a common transcript containing all messages that have been sent during the execution of a protocol instance. Every protocol can be transformed to provide this feature by requiring all parties to send all messages to all other parties and defining a fixed ordering on the sent messages – we consider an ordering of messages by the round they are sent, the index of the sender, and the receiver’s index in this sequence. If messages should be secret, each pair of parties executes a secure key exchange as part of the protocol instance and then encrypts messages with the established keys. Agreement is achieved by broadcasting signatures on the transcript, e.g., via signing the root of a Merkle tree over all message hashes as discussed in [FHKS21] and required in our transformations. Second, a *common public transcript of hashes* ensures that all parties hold a common

transcript containing the hashes of all messages sent during the execution of a protocol instance. This feature is achieved similar to the transcript of messages but parties only send message hashes to all parties that are not the intended receiver. Finally, the *private transcript* does not require any agreement on the transcript of a protocol instance.

Currently, all existing multi-party PVC protocols either provide a common public transcript of messages [DOS20, FHKS21] or a common public transcript of hashes [SSS21]. However, [DOS20] and [FHKS21] can be trivially adapted to provide just a common public transcript of hashes.

5 Building Blocks

In this section, we describe the building blocks for our financially backed covertly secure protocols. In the full version of this paper, we show security of the building blocks and that incorporating the building blocks into the PVC protocol does not affect the protocol’s security.

5.1 Internal State Commitments

To realize the judge in an efficient way, we want it to validate just a single protocol step instead of validating a whole instance. Existing PVC protocols prove misbehavior in a naive way by allowing parties to show that some other party P_j had an initial state $\text{state}_{(0)}^{(j)}$. Based on the initial state, the judge recomputes the whole protocol instance. In contrast to this, we incorporate a mechanism that allows parties to prove that P_j has been in state $\text{state}_{(k)}^{(j)}$ in a specific round k where misbehavior was detected. Then, the judge just needs to recompute a single step. To this end, we require that parties commit to each intermediate internal state during the execution of each semi-honest instance in a publicly verifiable way. In particular, in each round k of each semi-honest instance ℓ , each party P_i sends a hash of its internal state to all other parties using a collision-resistant hash function $H(\cdot)$, i.e., $H(\text{state}_{(\ell,k)}^{(i)})$. At the end of a protocol instance each party P_h creates a Merkle tree over all state hashes, i.e., $\text{sTree}_\ell := \text{MTree}(\{\text{hash}_{(\ell,k)}^{(i)}\}_{k \in [R], i \in [n]})$, and broadcasts a signature on the root of this tree, i.e., $\langle \text{MRoot}(\text{sTree}_\ell) \rangle_h$.

5.2 Signature Encoding

Our protocol incorporates signatures in order to provide evidence to the judge \mathcal{J} about the behavior of the parties. Without further countermeasures, an adversary can make use of signed data across multiple instances or rounds, e.g., she could claim that some message msg sent in round k has been sent in round k' using the signature received in round k . To prevent such an attack, we encode signed data by prefixing it with the corresponding indices before being signed. Merkle tree roots are prefixed with the instance index ℓ . Message hashes are

prefixed with ℓ , the round index k , the sender index i and the receiver index j . Initial state commitments ($\text{initData}_{(\ell,i)}^{\text{core}}$) are prefixed with ℓ and the index i of the party who's initial state the commitment refers to. The signature verification algorithm automatically checks for correct prefixing. The indices are derived from the super- and subscripts. If one index is not explicitly provided, e.g., in case only one instance is executed, the index is assumed to be 1.

5.3 Bisection of Trees

Our constructions make heavily use of Merkle trees to represent sets of data. This enables parties to efficiently prove that chunk of data is part of a set by providing a Merkle proof showing that the chunk is a leaf of the corresponding Merkle tree. In case two parties disagree about the data of a Merkle tree which should be identical, we use a bisection protocol Π_{BS} to narrow down the dispute to the first leaf of the tree on which they disagree. This helps a judging party to determine the lying party by just verifying a single data chunk in contrast to checking the whole data. The technique of bisecting was first used by Canetti et al. [CRR11] in the context of verifiable computing. Later, the technique was used in [KGC+18, TR19, EFS20].

The protocol is executed between a party P_b with input a tree mTree_b , a party P_m with input a tree mTree_m and a trusted judge \mathcal{J} announcing three public inputs: root^j , the root of mTree_j as claimed by P_j for $j \in \{b, m\}$, and width , the width of the trees, i.e., the number of leaves. The protocol returns the index z of the first leaf at which mTree_b and mTree_m differentiate, the leaf hash_z^m at position z of mTree_m , and the common leaf $\text{hash}_{(z-1)}$ at position $z - 1$. The latter is \perp if $z = 1$. Let $\text{node}(\text{mTree}, x, y)$ be the node of a tree mTree at position x of layer y – positions start with 1. The protocol is executed as follows:

Protocol Bisection Π_{BS}

1. \mathcal{J} initializes layer variable $y := 1$, position variable $x := 1$, last agreed hash $\text{hash}^a := \perp$, and $\text{depth} := \lceil \log_2(\text{width}) \rceil + 1$
2. All parties repeat this step while $y \leq \text{depth}$:
 - (a) Both P_j (for $j \in \{b, m\}$) send $\text{hash}^j := \text{node}(\text{mTree}_j, x, y)$ and $\sigma^j := \text{MProof}(\text{hash}^j, \text{mTree}_j)$ to \mathcal{J} .
 - (b) If $\text{MVerify}(\text{hash}^j, x, \text{root}^j, \sigma^j) = \text{false}$ (for $j \in \{b, m\}$), \mathcal{J} discards the message from P_j .
 - (c) If $y = \text{depth}$, \mathcal{J} keeps hash^b and hash^m and sets $y = y + 1$.
 - (d) If $y < \text{depth}$ and $\text{hash}^b = \text{hash}^m$, \mathcal{J} sets $x = (2 \cdot x) + 1$ and $y = y + 1$.
 - (e) If $y < \text{depth}$ and $\text{hash}^b \neq \text{hash}^m$, \mathcal{J} sets $x = (2 \cdot x) - 1$ and $y = y + 1$.
3. If $\text{hash}_b^b = \text{hash}_m^m$
 - \mathcal{J} sets $z := x + 1$ and $\text{hash}_{(z-1)} := \text{hash}_b^b$.
 - P_m sends $\text{hash}_z^m := \text{node}(\text{mTree}_m, z, \text{depth})$ and $\sigma := \text{MProof}(\text{hash}_z^m, \text{mTree}_m)$ to \mathcal{J} .
 - If $\text{MVerify}(\text{hash}_z^m, z, \text{root}^m, \sigma) = \text{false}$, \mathcal{J} discards. Otherwise \mathcal{J} stores hash_z^m .
4. If $\text{hash}_b^b \neq \text{hash}_m^m$
 - \mathcal{J} sets $z := x$ and $\text{hash}_z^m := \text{hash}_m^m$. If $z = 1$, \mathcal{J} sets $\text{hash}_{(z-1)} := \perp$, and the protocol jumps to step 5.

- P_m sends $\text{hash}_{(z-1)} := \text{node}(\text{mTree}_m, z - 1, \text{depth})$ and $\sigma := \text{MProof}(\text{hash}_{(z-1)}, \text{mTree}_m)$ to \mathcal{J} .
 - If $\text{MVerify}(\text{hash}_{(z-1)}, z - 1, \text{mTree}_m, \sigma) = \text{false}$, \mathcal{J} discards. Otherwise, \mathcal{J} keeps $\text{hash}_{(z-1)}$.
5. \mathcal{J} announces public outputs z , hash_z^m and $\text{hash}_{(z-1)}$.

6 Class 1: Input-Independent with Public Transcript

Our first transformation builds on input-independent PVC protocols where all parties possess a common public transcript of hashes (cf. Sect. 4.4) for each checked instance. Since the parties provide no input in these protocols, all parties can be opened. The set of input-independent protocols includes the important class of preprocessing protocols. In order to speed up MPC protocols, a common approach is to split the computation in an *offline* and an *online* phase. During the offline phase, precomputations are carried out to set up some correlated randomness. This phase does not require the actual inputs and can be executed continuously. In contrast, the online phase requires the private inputs of the parties and consumes the correlated randomness generated during the offline phase to speed up the execution. As the online performance is more time critical, the goal is to put as much work as possible into the offline phase. Prominent examples following this approach are the protocols of Damgård et al. [DPSZ12, DKL+13] and Wang et al. [WRK17a, WRK17b, YWZ20]. Input-independent PVC protocols with a public transcript can be obtained from semi-honest protocols using the input-independent compilers of Damgård et al. [DOS20] and Faust et al. [FHKS21].

In order to apply our construction to an input-independent PVC protocol, π^{PP} , we require π^{PP} to provide some features presented in Sect. 4 and to have incorporated some of the building blocks described in Sect. 5. First, we require the PVC protocol to be based on the cut-and-choose approach (cf. Sect. 4.1). Second, we require the actions of each party P_i in a protocol execution to be deterministically determined by a random seed (cf. Sect. 4.2). Third, we require that all parties learn the initial states of all other parties in the opened protocol instances (cf. Sect. 4.3). To this end, the parties receive signed data (e.g., a commitment and decommitment value) to derive the initial states of the other parties. Fourth, parties need to commit to their intermediate internal states during the protocol executions in a publicly verifiable way (cf. Sect. 5.1). Finally, all signed data match the encoded form specified in Sect. 5.2.

In order to achieve the public transcript of hashes and the commitments to the intermediate internal states, parties exchange additional data in each round. Formally, whenever some party P_h in round k of protocol instance ℓ transitions to a state $\text{state}_{(\ell, k)}^{(h)}$ with the outgoing messages $\{\text{msg}_{(\ell, k)}^{(h, i)}\}_{i \in [n] \setminus \{h\}}$, then it actually sends the following to P_i :

$$(\text{msg}_{(\ell, k)}^{(h, i)}, \{\text{hash}_{(\ell, k)}^{(h, j)} := H(\text{msg}_{(\ell, k)}^{(h, j)})\}_{j \in [n] \setminus \{h, i\}}, \text{hash}_{(\ell, k)}^{(h)} := H(\text{state}_{(\ell, k)}^{(h)}))$$

Let O denote the set of opened instances. We summarize the aforementioned requirements by specifying the data that the view of any honest party P_h includes. It contains signed data to derive the initial state of all parties for the opened instances (1a), a Merkle tree over the hashes of all messages exchanged within a single instance for all instances (1b), a Merkle tree over the hashes of all intermediate internal states of a single instance for all instances (1c), and signatures from each party over the roots of the message and state trees (1d):

$$\{\langle \text{initData}_{(i,\ell)}^{\text{core}} \rangle_i, \text{initData}_{(i,\ell)}^{\text{aux}}\}_{\ell \in O, i \in [n]}, \quad (1a)$$

$$\{\text{mTree}_\ell\}_{\ell \in [t]} := \{\text{MTree}(\{\text{hash}_{(\ell,k)}^{(i,j)}\}_{k \in [R], i \in [n], j \neq i})\}_{\ell \in [t]}, \quad (1b)$$

$$\{\text{sTree}_\ell\}_{\ell \in [t]} := \{\text{MTree}(\{\text{hash}_{(\ell,k)}^{(i)}\}_{k \in [R], i \in [n]})\}_{\ell \in [t]} \quad (1c)$$

$$\{\langle \text{MRoot}(\text{mTree}_\ell) \rangle_i\}_{i \in [n], \ell \in [t]} \text{ and } \{\langle \text{MRoot}(\text{sTree}_\ell) \rangle_i\}_{i \in [n], \ell \in [t]}. \quad (1d)$$

We next define the blame algorithm that takes the specified view as input and continue with the description of the punishment protocol afterwards.

The Blame Algorithm. At the end of protocol π^{PP} , all parties execute the blame algorithm Blame^{PP} to generate a certificate cert . The resulting certificate is broadcasted and the honest party finishes the execution of π^{PP} by outputting cert . The certificate is generated as follows:

Algorithm Blame^{PP}

1. P_h runs $\text{state}_{(\ell,0)}^{(i)} = \text{derivInit}(\text{initData}_{(i,\ell)}^{\text{core}}, \text{initData}_{(i,\ell)}^{\text{aux}})$ for each $i \in [n], \ell \in O$. Let \mathcal{B} be the set of all tuples $(\ell, 0, m, 0)$ such that $\text{state}_{(\ell,0)}^{(m)} = \text{bad}$. If $\mathcal{B} \neq \emptyset$, goto step 4.
2. P_h emulates for each $\ell \in O$ the protocol executions on input the initial states from all parties to obtain the expected messages and the expected intermediate states of all parties, i.e., $(\{\text{msg}_{(\ell,k)}^{(i,j)}\}_{k \in [R], i \in [n], j \neq i}, \{\text{state}_{(\ell,k)}^{(i)}\}_{k,i,j}) := \text{emulate}^{\text{full}}(\{\text{state}_{(\ell,0)}^{(i)}\}_{i \in [n]})$.
3. Let \mathcal{B} be the set of all tuples (ℓ, k, m, i) such that $H(\text{msg}_{(\ell,k)}^{(m,i)}) \neq \text{hash}_{(\ell,k)}^{(m,i)}$ or $H(\text{state}_{(\ell,k)}^{(m)}) \neq \text{hash}_{(\ell,k)}^{(m)}$ – where $\text{hash}_{(\ell,k)}^{(m,i)}$ and $\text{hash}_{(\ell,k)}^{(m)}$ are extracted from mTree_ℓ or sTree_ℓ respectively. In case of an incorrect state hash, set $i = 0$.
4. If $\mathcal{B} = \emptyset$ P_h outputs $\text{cert} := \perp$. Otherwise, P_h picks the tuple (ℓ, k, m, i) from \mathcal{B} with the smallest ℓ, k, m, i in this sequence, sets $k' := k - 1$ and defines variables as follows – variables that are not explicitly defined are set to \perp .

(Always): $ids := (\ell, k, m, i)$
 $initData := (\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m)}^{\text{aux}})$
 $root^{\text{state}} := \langle \text{MRoot}(\text{sTree}_\ell) \rangle_m$
 $root^{\text{msg}} := \langle \text{MRoot}(\text{mTree}_\ell) \rangle_m$
(If $k > 0$): $state_{out} := (\text{hash}_{(\ell, k)}^{(m)}, \text{MProof}(\text{hash}_{(\ell, k)}^{(m)}, \text{sTree}_\ell))$
 $msg_{out} := (\text{hash}_{(\ell, k)}^{(m, i)}, \text{MProof}(\text{hash}_{(\ell, k)}^{(m, i)}, \text{mTree}_\ell))$
(If $k > 1$): $state_{in} := (\text{state}_{(\ell, k')}^{(m)}, \text{MProof}(H(\text{state}_{(\ell, k')}^{(m)}), \text{sTree}_\ell))$
 $\mathcal{M}_{in} := \{(\text{msg}_{(\ell, k')}^{(j, m)}, \text{MProof}(H(\text{msg}_{(\ell, k')}^{(j, m)}), \text{mTree}_\ell))\}_{j \in [n]}$

5. Output $\text{cert} := (ids, \text{initData}, \text{root}^{\text{state}}, \text{root}^{\text{msg}}, \text{state}_{in}, \mathcal{M}_{in}, \text{state}_{out}, \text{msg}_{out})$.

The Punishment Protocol. Each party P_i (for $i \in [n]$) checks if $\text{cert} \neq \perp$. If this is the case, P_i sends cert to \mathcal{J}^{PP} . Otherwise, P_i waits till time T to receive her deposit back. Timeout T is set such that the parties have sufficient time to submit a certificate after the execution of π^{PP} and Blame^{PP} . The judge \mathcal{J}^{PP} is described in the following. The validation algorithms wrongMsg and wrongState and the algorithm getIndex can be found in the full version of this paper. We stress that the validation algorithms wrongMsg and wrongState don't need to recompute a whole protocol execution but only a single step. Therefore, \mathcal{J}^{PP} is very efficient and can, for instance, be realized via a smart contract. To be more precise, the judge is execution without any interaction and runs in computation complexity linear in the protocol complexity. By allowing logarithmic interactions between the judge and the parties, we can further reduce the computation complexity to logarithmic in the protocol complexity. This can be achieved by applying the efficiency improvement described in the full version of this paper.

Judge \mathcal{J}^{PP}

Initialization: The judge has access to public variables n, t, T and the set of parties $\{P_i\}_{i \in [n]}$. Further, it maintains a set cheaters initially set to \emptyset . Prior to the execution of π^{PP} , \mathcal{J}^{PP} has received d coins from each party P_i .

Proof verification: Wait until time T_1 to receive $(\ell, k, m, i), \text{initData}, \langle \text{root}_{(\ell)}^{\text{state}} \rangle_m, \langle \text{root}_{(\ell)}^{\text{msg}} \rangle_m, \text{state}_{in}, \mathcal{M}_{in}, \text{state}_{out}, (\text{hash}, \sigma)$ and do:

1. If $P_m \in \text{cheaters}$, abort.
2. Parse initData to $(\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m)}^{\text{aux}})$ and set $\text{state}_0 = \text{derivelnit}(\text{initData}_{(\ell, m)}^{\text{core}}, \text{initData}_{(\ell, m)}^{\text{aux}})$. If $\text{Verify}(\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m) = \text{false}$ or $\text{state}_0 = \perp$, abort. If $\text{state}_0 = \text{bad}$, add P_m to cheaters and stop.
3. If $\text{Verify}(\langle \text{root}_{(\ell)}^{\text{state}} \rangle_m) = \text{false}$ or $\text{Verify}(\langle \text{root}_{(\ell)}^{\text{msg}} \rangle_m) = \text{false}$, abort.
4. If $i = 0$ and $\text{wrongState}(\text{state}_0, \text{state}_{in}, \text{state}_{out}, \mathcal{M}_{in}, \text{root}_{(\ell)}^{\text{state}}, \text{root}_{(\ell)}^{\text{msg}}, \ell, k, m) = \text{true}$, add P_m to cheaters .

5. If $i > 0$, $MVerify(\text{hash}, \text{getIndex}(k, m, i), \text{root}_{(\ell)}^{\text{msg}}, \sigma) = \text{true}$ and $\text{wrongMsg}(\text{state}_0, \text{state}_{in}, \text{hash}, \mathcal{M}_{in}, \text{root}_{(\ell)}^{\text{state}}, \text{root}_{(\ell)}^{\text{msg}}, \ell, m, k, i) = \text{true}$, add P_m to cheaters.

Timeout: At time T_1 , send d coins to each party $P_i \notin \text{cheaters}$.

6.1 Security

Theorem 1. *Let $(\pi^{\text{PP}}, \cdot, \cdot)$ be an n -party publicly verifiable covert protocol computing function f with deterrence factor ϵ satisfying the view requirements stated in Eq. (1a)–(1d). Further, let the signature scheme (Generate, Sign, Verify) be existentially unforgeable under chosen-message attacks, the Merkle tree satisfies the binding property and the hash function H be collision resistant. Then the protocol π^{PP} together with algorithm Blame^{PP} , protocol $\text{Punish}^{\text{PP}}$ and judge \mathcal{J}^{PP} satisfies financially backed covert security with deterrence factor ϵ according to Definition 1.*

We formally prove Theorem 1 in the full version of this paper.

7 Class 3: Input-Independent with Private Transcript

At the time of writing, there exists no PVC protocol without public transcript that could be directly transformed into an FBC protocol. Moreover, it is not clear, if it is possible to construct a PVC protocol without a public transcript. Instead, we present a transformation from an input-independent PVC protocol with public transcript into an FBC protocol without any form of common public transcript. As in our first transformation, we start with an input-independent PVC protocol π_3^{PVC} that is based on cut-and-choose where parties share a common public transcript. Due to the input-independence, all parties of the checked instances can be opened. However, unlike our first transformation, which utilizes the public transcript, we remove this feature from the PVC protocol as part of the transformation. We denote the protocol that results by removing the public transcript feature from π_3^{PVC} by π_3 . Without having a public transcript, the punishment protocol becomes interactive and more complicated. Intuitively, without a public transcript it is impossible to immediately decide if a message that deviates from the emulation is maliciously generated or is invalid because of a received invalid messages. Note that we still have a common public tree of internal state hashes in our exposition. However, the necessity of this tree can also be removed by applying the techniques presented here that allow us to remove the common transcript.

In order to apply our construction to a protocol π_3 , we require almost the same features of π_3 as demanded in our first transformation (cf. Sect. 6). For the sake of exposition, we outline the required features here again and point out the differences. First, we require π_3 to be based on the cut-and-choose approach (cf. Sect. 4.1). Second, we require the actions of each party P_i in a semi-honest

instance execution to be deterministically determined by a random seed (cf. Sect. 4.2). Third, we require that all parties learn the initial states of all other parties in the opened protocol instances (cf. Sect. 4.3). To this end, the parties receive signed data (e.g., a commitment and decommitment value) to derive the initial states of the other parties. Fourth, parties need to commit to their intermediate internal states during the protocol executions in a publicly verifiable way (cf. Sect. 5.1). Finally, all signed data match the encoded form specified in Sect. 5.2.

In contrast to the transformation in Sect. 6 we no longer require from protocol π_3 that the parties send all messages or message hashes to all other parties. Formally, whenever some party P_h in round k of protocol instance ℓ transitions to a state $\mathbf{state}_{(\ell,k)}^{(h)}$ with the outgoing messages $\{\mathbf{msg}_{(\ell,k)}^{(h,i)}\}_{i \in [n] \setminus \{h\}}$, then it actually sends the following to P_i :

$$\langle \langle \mathbf{msg}_{(\ell,k)}^{(h,i)} \rangle_h, \mathbf{hash}_{(\ell,k)}^{(h)} := H(\mathbf{state}_{(\ell,k)}^{(h)}) \rangle$$

Let O be the set of opened instances. We summarize the aforementioned requirements by specifying the data that the view of any honest party P_h after the execution of π_3 includes. The view contains data to derive the initial state of all parties which is signed by each party for each party and every opened instance, i.e.,

$$\{ \langle \langle \mathbf{initData}_{(i,\ell)}^{\text{core}} \rangle_j, \mathbf{initData}_{(i,\ell)}^{\text{aux}} \rangle \}_{\ell \in O, i \in [n], j \in [n]}, \quad (2a)$$

a Merkle tree over the hashes of all intermediate internal states of a single instance for all instances, i.e.,

$$\{ \mathbf{sTree}_\ell \}_{\ell \in [t]} := \{ \mathbf{MTree}(\{ \mathbf{hash}_{(\ell,k)}^{(i)} \}_{k \in [R], i \in [n]}) \}_{\ell \in [t]}, \quad (2b)$$

signatures from each party over the roots of the state trees, i.e.,

$$\{ \langle \mathbf{MRoot}(\mathbf{sTree}_\ell) \rangle_i \}_{i \in [n], \ell \in [t]} \quad (2c)$$

and the signed incoming message, i.e.,

$$\mathcal{M} := \{ \langle \mathbf{msg}_{(\ell,k)}^{(i,h)} \rangle_i \}_{\ell \in [t], k \in [R], i \in [n] \setminus \{h\}}. \quad (2d)$$

The Blame Algorithm. At the end of protocol π_3 , all parties first execute an evidence algorithm **Evidence** to generate partial certificates \mathbf{cert}' . The partial certificate is a candidate to be used for the punishment protocol and is broadcasted to all other parties as part of π_3 . In case the honest party detects cheating in several occurrences, the party picks the occurrence with the smallest indices (ℓ, k, m, i) (in this sequence). The algorithm to generate partial certificates **Evidence** is formally described as follows:

Algorithm Evidence

1. P_h runs $\text{state}_{(\ell,0)}^{(i)} = \text{derivelnit}(\text{initData}_{(i,\ell)}^{\text{core}}, \text{initData}_{(i,\ell)}^{\text{aux}})$ for each $i \in [n], \ell \in O$. Let \mathcal{B} be the set of all tuples $(\ell, 0, m, 0)$ such that $\text{state}_{(\ell,0)}^{(m)} = \text{bad}$. If $\mathcal{B} \neq \emptyset$, goto step 4.
2. P_h emulates for each $\ell \in O$ the protocol executions on input the initial states from all parties to obtain the expected messages and the expected intermediate states of all parties, i.e., $(\{\widetilde{\text{msg}}_{(\ell,k)}^{(i,j)}\}_{k \in [R], i \in [n], j \neq i}, \{\text{state}_{(\ell,k)}^{(i)}\}_{k,i,j}) := \text{emulate}^{\text{full}}(\{\text{state}_{(\ell,0)}^{(i)}\}_{i \in [n]})$.
3. Let \mathcal{B} be the set of all tuples (ℓ, k, m, h) such that $\text{msg}_{(\ell,k)}^{(m,h)} \neq \widetilde{\text{msg}}_{(\ell,k)}^{(m,h)}$ or $H(\text{state}_{(\ell,k)}^{(m)}) \neq \text{hash}_{(\ell,k)}^{(m)}$ – where $\text{msg}_{(\ell,k)}^{(m,h)}$ and $\text{hash}_{(\ell,k)}^{(m)}$ are taken from \mathcal{M} or sTree_ℓ respectively. In case of an invalid state, set $h = 0$.
4. Pick the tuple (ℓ, k, m, i) from \mathcal{B} with the smallest ℓ, k, m, i in this sequence. If $k > 0$ set $\text{msg}_{\text{out}} := \langle \text{msg}_{(\ell,k)}^{(m,i)} \rangle_m$. Otherwise, set $\text{msg}_{\text{out}} := \perp$.
5. Output partial certificate $(ids, \text{msg}_{\text{out}})$.

Since π_3 does not contain a public transcript of messages, parties can only validate their own incoming message instead of all messages as done in previous approaches. Hence, it can happen that different honest parties generate and broadcast different partial certificates. Therefore, all parties validate the incoming certificates, discard invalid ones and pick the partial certificate cert' with the smallest indices (ℓ, k, m, i) (in this sequence) as their own. If no partial certificate has been received or created, parties set $\text{cert}' := \perp$.

Finally, each honest party executes the blame algorithm Blame^{SP} to create the full certificate that is used for both, blaming a malicious party and defending against incorrect accusations. As in this scenario the punishment protocol requires input of accused honest parties, the blame algorithm returns a certificate even if no malicious behavior has been detected, i.e., if $\text{cert}' = \perp$. The final certificate is generated by appending following data from the view to the certificate: $\{(\langle \text{initData}_{(i,\ell)}^{\text{core}} \rangle_j, \text{initData}_{(i,\ell)}^{\text{aux}})\}_{\ell \in O, i \in [n], j \in [n]}$ (cf. Eq. 2a), $\{\text{sTree}_\ell\}_{\ell \in [t]}$ (cf. Eq. 2b), and $\{\langle \text{MRoot}(\text{sTree}_\ell) \rangle_i\}_{i \in [n], \ell \in [t]}$ (cf. Eq. 2c). All the appended data is public and does not really need to be broadcasted. However, in order to match the formal specification, all parties broadcast their whole certificate. If $\text{cert}' \neq \perp$, the honest party outputs in addition to the certificate corrupted_m .

To ease the specification of the punishment protocol in which parties derive further data from the certificates, we define an additional algorithm mesHistory that uses the messages obtained during the emulation $(\widetilde{\text{msg}})^2$ to compute the message history up to a specific round k' (inclusively) of instance ℓ . We structure the message history in two layers. For each round $k^* < k'$, parties create a Merkle tree of all messages emulated in this round. These trees constitute the bottom layer. On the top layer, parties create a Merkle tree over the roots of the bottom layer trees. This enables parties to agree on all messages of one round making

² Formally, parties need to re-execute the emulation, as we do not allow them to use any data not included in the certificate.

it easier to submit Merkle proofs for messages sent in this round. The message history is composed of the following variables:

$$\begin{aligned} \{\text{mTree}_{k^*}^{\text{round}}\}_{k^* \in [k']} &:= \{\text{MTree}(\{H(\widetilde{\text{msg}}_{(\ell, k^*)}^{(i, j)})\}_{i \in [n], j \neq i})\}_{k^* \in [k']} \\ \text{mTree}_{k'} &:= \text{MTree}(\{\text{MRoot}(\text{mTree}_{k^*}^{\text{round}})\}_{k^* \in [k']}) \\ \text{root}_{k'}^{\text{msg}} &:= \text{MRoot}(\text{mTree}) \end{aligned}$$

Additionally, if $\text{cert}' \neq \perp$, parties compute the following:

$$\begin{aligned} \text{(Always): } \text{initData} &:= (\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m)}^{\text{aux}}) \\ \text{root}^{\text{state}} &:= \langle \text{MRoot}(\text{sTree}_\ell) \rangle_m \\ \text{(If } k > 0\text{): } \text{state}_{\text{out}} &:= (\text{hash}_{(\ell, k)}^{(m)}, \text{MProof}(\text{hash}_{(\ell, k)}^{(m)}, \text{sTree}_\ell)) \\ \text{(If } k > 1\text{): } \text{state}_{\text{in}} &:= (\text{state}_{(\ell, k')}^{(m)}, \text{MProof}(H(\text{state}_{(\ell, k')}^{(m)}), \text{sTree}_\ell)) \\ &(\{\text{mTree}_{k^*}^{\text{round}}\}_{k^* \in [k']}, \text{mTree}_{k'}, \text{root}_{k'}^{\text{msg}}) := \text{mesHistory}(k', \ell) \\ \sigma_{k'} &:= \text{MProof}(\text{MRoot}(\text{mTree}_{k'}^{\text{round}}), \text{mTree}_{k'}) \\ \mathcal{M}_{\text{in}} &:= \{(\widetilde{\text{msg}}_{(\ell, k')}^{(j, m)}, \text{MProof}(H(\widetilde{\text{msg}}_{(\ell, k')}^{(j, m)}), \text{mTree}_{k'}^{\text{round}}))\}_{j \in [n]} \end{aligned}$$

The Punishment Protocol. The main difficulty of constructing a punishment protocol $\text{Punish}^{\text{SP}}$ for this scenario is that there is no publicly verifiable evidence about messages like a common transcript used in the previous transformations. Hence, incoming messages required for the computation of a particular protocol step cannot be validated directly. Instead, the actions of all parties need to be validated against the emulated actions based on the initial states. This leads to the problem that deviations from the protocol can cause later messages of other honest parties to deviate from the emulated ones as well. Therefore, it is important that the judge disputes the earliest occurrence of misbehavior.

We divide the punishment protocol $\text{Punish}^{\text{SP}}$ into three phases. First, the judge determines the earliest accusation of misbehavior. To this end, if $\text{cert} \neq \perp$ all parties start by sending tuple ids from cert to \mathcal{J}^{SP} and the judge selects the tuple with the smallest indices (ℓ, k, m, i) . This mechanism ensures that either the first malicious message or malicious state hash received by an honest party is disputed or the adversary blames some party at an earlier point. To look ahead, if the adversary blames an honest party at an earlier point, the punishment will not be successful and the malicious blamer will be punished for submitting an invalid accusation. If the adversary blames another malicious party, either one of them will be punished. This mechanism ensures that if an honest party submits an accusation, a malicious party will be punished, even if it is not the honest party's accusation that is disputed.

If there has not been any accusation submitted in the first phase, \mathcal{J}^{SP} reimburses all parties. Otherwise, \mathcal{J}^{SP} defines a blamer P_b , the party that has submitted the earliest accusation, and an accused party P_m . P_b either accuses misbehavior in the initial state, the first round, or in some later round. For the

former two, misbehavior can be proven in a straightforward way, similar to our first construction. For the latter, P_b is supposed to submit a proof containing the hash of a tree of the message history up to the disputed round k . P_m can accept or decline the message history depending on whether the tree corresponds to the one emulated by P_m or not. If the tree is accepted, the certificate can be validated as in previous scenarios, with the only difference that incoming messages are validated with respect to the submitted message history tree instead of the common public transcript. In case any party does not respond in time, this party is considered maliciously and is financially punished.

If the message history is declined, the protocol transitions to the third phase. Parties P_b and P_m together with \mathcal{J}^{SP} execute a bisection search in the message history tree to find the first message they disagree on (cf. Sect. 5.3). By definition they agree on all messages before the disputed one – we call these messages the *agreed sub-tree*. At this step, \mathcal{J}^{SP} can validate the disputed message of the history tree (not the one disputed in the beginning) the same way as done in previous constructions with the only difference that incoming messages are validated with respect to the agreed sub-tree.

The number of interactions is logarithmic while the computation complexity of the judge is linear in the protocol complexity. We can further reduce the computation complexity to be logarithmic in the protocol complexity while still having logarithmic interactions using the efficiency improvements described in the full version of this paper. The judge is defined as follows:

Protocol Punish^{SP}

Phase 1: Determine earliest accusation

1. If $\text{cert} \neq \perp$, P_h sends $\text{ids} := (\ell, k, m, i)$ taken from cert to \mathcal{J}^{SP} which stores (ℓ, k, m, i, h) .
2. \mathcal{J}^{SP} waits till time T to receive message (ℓ, k, m, i) from parties P_b for $b \in [n]$. If no accusations have been received, \mathcal{J}^{SP} sends d coins to each party at time T . Otherwise, \mathcal{J}^{SP} picks the *smallest* tuple (ℓ, k, m, i, b) (ordered in this sequence), sets $k' := k - 1$ and continues with Phase 2.

Timeout: If its P_j 's turn for $j \in \{b, m\}$ and P_j does not respond with a valid message, i.e., one that is not discarded, in time, P_j is considered malicious and \mathcal{J}^{SP} terminates by sending d coins to all parties but P_j .

Phase 2: First evidence

3. If $k < 2$, P_b sends $(\text{initData}, \text{root}^{\text{state}}, \text{state}_{\text{out}}, \langle \text{msg}_{(\ell, k)}^{(m, i)} \rangle_m)$ taken from cert to \mathcal{J}^{SP}
 - (a) \mathcal{J}^{SP} parses initData to $(\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m)}^{\text{aux}})$ and sets $\text{state}_0 = \text{derivelnit}(\text{initData}_{(\ell, m)}^{\text{core}}, \text{initData}_{(\ell, m)}^{\text{aux}})$. If $\text{Verify}(\langle \text{initData}_{(\ell, m)}^{\text{core}} \rangle_m) = \text{false}$ or $\text{state}_0 = \perp$, \mathcal{J}^{SP} discards. If $\text{state}_0 = \text{bad}$, \mathcal{J}^{SP} terminates by sending d coins to all parties but P_m .
 - (b) If $\text{Verify}(\langle \text{root}_{(\ell)}^{\text{state}} \rangle_m) = \text{false}$, \mathcal{J}^{SP} discards.
 - (c) If $i = 0$ and $\text{wrongState}(\text{state}_0, \perp, \text{state}_{\text{out}}, \emptyset, \text{root}_{(\ell)}^{\text{state}}, \perp, \ell, k, m) = \text{false}$, \mathcal{J}^{SP} discards.
 - (d) If $i > 0$, $\text{Verify}(\langle \text{msg}_{(\ell, k)}^{(m, i)} \rangle_m) = \text{false}$ or $\text{wrongMsg}(\text{state}_0, \perp, H(\text{msg}_{(\ell, k)}^{(m, i)}), \emptyset, \text{root}_{(\ell)}^{\text{state}}, \perp, \ell, m, k, i) = \text{false}$, \mathcal{J}^{SP} discards.

- (e) \mathcal{J}^{SP} terminates by sending d coins to all parties but P_m .
4. Otherwise, P_b sends $(\text{root}^{\text{state}}, \text{state}_{in}, \text{state}_{out}, \langle \text{root}_{(\ell)}^{\text{state}} \rangle_m, \text{root}^{\text{msg}}, \text{root}_{k'}^{\text{round}}, \sigma_{k'}, \mathcal{M}_{in}, \text{msg}_{out})$ taken from cert to \mathcal{J}^{SP} .
- (a) P_m executes $\text{mesHistory}(k-1, \ell)$. Let $\widetilde{\text{root}}^{\text{msg}}$ be the root of the emulated message history tree. If $\text{root}^{\text{msg}} \neq \widetilde{\text{root}}^{\text{msg}}$ P_m sends $\widetilde{\text{root}}^{\text{msg}}$ to \mathcal{J}^{SP} . Otherwise, P_m sends \perp .
- (b) If $\widetilde{\text{root}}^{\text{msg}}$ received by P_m does not equal \perp , \mathcal{J}^{SP} jumps to phase 3.
- (c) \mathcal{J}^{SP} checks that $\text{Verify}(\langle \text{root}_{(\ell)}^{\text{state}} \rangle_m) = \text{true}$ and $\text{MVerify}(\text{root}_{k'}^{\text{round}}, k', \text{root}^{\text{msg}}, \sigma_{k'}) = \text{true}$ and discards otherwise.
- (d) If $i = 0$ and $\text{wrongState}(\perp, \text{state}_{in}, \text{state}_{out}, \mathcal{M}_{in}, \text{root}_{(\ell)}^{\text{state}}, \text{root}_{k'}^{\text{round}}, \ell, k, m) = \text{false}$, \mathcal{J}^{SP} discards.
- (e) If $i > 0$, $\text{Verify}(\langle \text{msg}_{(\ell, k)}^{(m, i)} \rangle_m) = \text{false}$ or $\text{wrongMsg}(\text{state}_0, \text{state}_{in}, H(\text{msg}_{(\ell, k)}^{(m, i)}), \mathcal{M}_{in}, \text{root}_{(\ell)}^{\text{state}}, \text{root}_{k'}^{\text{round}}, \ell, m, k, i) = \text{false}$, \mathcal{J}^{SP} discards.
- (f) \mathcal{J}^{SP} terminates by sending d coins to all parties but P_m .

Phase 3: Dispute the message tree

5. Parties P_b , P_m and \mathcal{J}^{SP} run bisection sub-protocol Π_{BS} on the top-level tree. P_b 's input is the tree with root root^{msg} ; P_m 's the one with root $\widetilde{\text{root}}^{\text{msg}}$. \mathcal{J}^{SP} announces public inputs root^{msg} and width of root^{msg} , $\text{width} := k'$. The output is the first round they disagree on k_2 , the agreed hash $\text{root}_{k_2'}^{\text{round}}$ of leaf with index $k_2' := k_2 - 1$ and the hash $\text{root}_{(b, k_2)}^{\text{round}}$ of leaf with index k_2 as claimed by P_m .
6. Parties P_m , P_b and \mathcal{J}^{SP} run bisection sub-protocol Π_{BS} on the low-level tree. Both, P_m and P_b take as input $\text{mTree}_{k_2}^{\text{round}}$ from their certificate. \mathcal{J}^{SP} announces public inputs $\text{root}_{(b, k_2)}^{\text{round}}$ and the width of the low level tree $\text{width}'_n \times (n-1)$. The output is the index x of the first message they disagree on and the hash of this message hash_x as claimed by P_m . The index of the sender of the disputed message is $m_2 := \lceil \frac{x}{n-1} \rceil$ and the index of the receiver $i_2 = x \bmod (n-1)$ if $m_2 > (x \bmod (n-1))$ and $i_2 := (x \bmod (n-1)) + 1$ otherwise.
7. Party P_b define variables as follows – variables that are not explicitly defined are set to \perp .

$$\begin{aligned} \text{(Always): } \text{initData}^2 &:= (\langle \text{initData}_{(\ell, m_2)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m_2)}^{\text{aux}}) \\ \text{root}^{\text{state}} &:= \langle \text{MRoot}(\text{sTree}_{\ell}) \rangle_m \end{aligned}$$

$$\begin{aligned} \text{(If } k_2 > 1\text{): } \text{state}_{in}^2 &:= (\text{state}_{(\ell, k_2')}^{(m_2)}, \text{MProof}(H(\text{state}_{(\ell, k_2')}^{(m_2)}), \text{sTree}_{\ell})) \\ \mathcal{M}_{in}^2 &:= \{(\text{msg}_{(\ell, k_2')}^{(j, m_2)}, \text{MProof}(H(\text{msg}_{(\ell, k_2')}^{(j, m_2)}), \text{mTree}_{k_2}^{\text{round}}))\}_{j \in [n]} \end{aligned}$$

and sends $(\text{initData}^2, \langle \text{MRoot}(\text{sTree}_{\ell}) \rangle_m, \text{state}_{in}^2, \mathcal{M}_{in}^2)$ to \mathcal{J}^{SP} .

8. \mathcal{J}^{SP} parses initData^2 to $(\langle \text{initData}_{(\ell, m_2)}^{\text{core}} \rangle_m, \text{initData}_{(\ell, m_2)}^{\text{aux}})$ and sets $\text{state}_{(0)}^{(m_2)} := \text{derivelnit}(\text{initData}_{(\ell, m_2)}^{\text{core}}, \text{initData}_{(\ell, m_2)}^{\text{aux}})$. If $\text{Verify}(\langle \text{root}_{(\ell)}^{\text{state}} \rangle_m) = \text{false}$, $\text{Verify}(\langle \text{initData}_{(\ell, m_2)}^{\text{core}} \rangle_m) = \text{false}$ or $\text{state}_{(0)}^{(m_2)} \in \{\perp, \text{bad}\}$, \mathcal{J}^{SP} discards.
9. If $\text{wrongMsg}(\text{state}_{(0)}^{(m_2)}, \text{state}_{in}^2, \text{hash}_x, \mathcal{M}_{in}^2, \text{root}_{(\ell)}^{\text{state}}, \text{root}_{k_2'}^{\text{round}}, \ell, m_2, k_2, i_2) = \text{false}$, \mathcal{J}^{SP} discards.
10. \mathcal{J}^{SP} terminates by sending d coins to all parties but P_m .

7.1 Security

Theorem 2. *Let $(\pi_3^{\text{PVC}}, \text{Blame}^{\text{PVC}}, \text{Judge}^{\text{PVC}})$ be an n -party publicly verifiable covert protocol computing function f with deterrence factor ϵ satisfying the view requirements stated in Eq. (2). Further, π_3^{PVC} generates a common public transcript of hashes that is only used for $\text{Blame}^{\text{PVC}}$ and $\text{Judge}^{\text{PVC}}$. Let π_3 be a protocol that is equal to π_3^{PVC} but does not generate a common transcript and instead of calling $\text{Blame}^{\text{PVC}}$ executes the blame procedure explained above (including execution of Evidence and Punish^{SP}). Further, let the signature scheme (Generate, Sign, Verify) be existentially unforgeable under chosen-message attacks, the Merkle tree satisfies the binding property, the hash function H be collision resistant and the bisection protocol Π_{BS} be correct. Then, the protocol π_3 , together with algorithm Blame^{SP} , protocol Punish^{SP} and judge \mathcal{J}^{SP} satisfies financially backed covert security with deterrence factor ϵ according to Definition 1.*

We formally prove Theorem 2 in the full version of this paper.

8 Evaluation

In order to evaluate the practicability of our protocols, i.e., to show that the judging party can be realized efficiently via a smart contract, we implemented the judge of our third transformation (cf. Sect. 7) for the Ethereum blockchain and measured the associated execution costs. We focus on the third setting, the verification of protocols with a private transcript, since we expect this scenario to be the most expensive one due to the interactive punishment procedure. Further, we have extended the transformation such that the protocol does not require a public transcript of state hashes.

Our implementation includes the efficiency features described in the full version of this paper. In particular, we model the calculation of each round’s and party’s `computeRound` function as an arithmetic circuit and compress disputed calculations and messages using Merkle trees. The latter are divided into 32-byte chunks which constitute the leaf of the Merkle tree. The judge only needs to validate either the computation of a single arithmetic gate or the correctness of a single message chunk of a sent or received message together with the corresponding Merkle tree proofs. The proofs are logarithmic in the size of the computation resp. the size of a message. Messages are validated by defining a mapping from each chunk to a gate in the corresponding `computeRound` function.

In order to avoid redundant deployment costs, we apply a pattern that allows us to deploy the contract code just once and for all and create new independent instances of our FBC protocol without deploying further code. When starting a new protocol instance, parties register the instance at the existing contract which occupies the storage for the variables required by the new instance, e.g., the set of involved parties. Further, we implement the judge to be agnostic to the particular semi-honest protocol executed by the parties – recall that our FBC protocol wraps around a semi-honest protocol that is subject to the cut-and-choose technique. Every instance registered at the judge can involve a different

Table 1. Costs for deployment, instance registration and optimistic execution.

Protocol steps	n	Cost	
		Gas	USD
Deployment		4 775 k	639.91
New instance	2	287 k	38.41
New instance	3	308 k	41.30
New instance	5	351 k	47.05
New instance	10	458 k	61.43
Honest execution	2	178 k	23.92
Honest execution	3	224 k	30.07
Honest execution	5	316 k	42.38
Honest execution	10	546 k	73.14

Gates: Number of gates in the circuit of each `computeRound` function.

Chunks: Number of chunks in each message.

R: Number of communication rounds.

n: Number of parties.

Table 2. Worst-case execution costs.

Gates	Chunks	R	n	Cost	
				Gas	USD
10	10	10	3	1 780 k	238.58
1 000	10	10	3	2 412 k	323.25
1 M	10	10	3	3 512 k	470.55
1 B	10	10	3	4 782 k	640.75
1 T	10	10	3	6 182 k	828.35
10	10	10	3	1 785 k	239.14
100	100	10	3	2 086 k	279.61
1 000	1 000	10	3	2 422 k	324.55
100	10	10	3	2 081 k	278.91
100	10	10	4	2 223 k	297.86
100	10	10	7	2 442 k	327.29
100	10	10	10	2 659 k	356.34
100	10	10	50	4 764 k	638.35
100	10	3	3	1 878 k	251.65
100	10	10	3	2 074 k	277.88
100	10	100	3	2 403 k	322.04
100	10	1 000	3	2 834 k	379.79

number of parties and define its own semi-honest protocol. This means that the same judge contract can be used for whatever semi-honest protocol our FBC protocol instance is based on, e.g., for both the generation of Beaver triples and garbled circuits. Parties simply define for each involved party and each round the `computeRound` function as a set of gates, aggregate all gates into a Merkle tree and submit the tree’s root upon instance registration.

We perform all measurements on a local test environment. We setup the local Ethereum blockchain with *Ganache* (core version 2.13.2) on the latest supported hard fork, Muir Glacier. The contract is compiled to EVM byte code with *solc* (version 0.8.1, optimized on 20 runs). As common, we measure the efficiency of the smart contracts via its gas consumption – this metric directly translates to execution costs. Further, we estimate USD costs based on the prices (gas to ETH and ETH to USD) on Aug. 20, 2021 [Eth21, Coi21]. For comparison, a simple Ether transfer costs 21,000 gas resp. 2,81 USD.

In Table 1, we display the costs of the deployment, the registration of a new instance and the optimistic execution without any disputes. The costs of these steps only depend on the number of parties. In Table 2, we display the worst-case costs of a protocol execution for different protocol parameters, i.e., complexity of the `computeRound` functions, message size, communication rounds and number of parties. In order to determine the worst-case costs, we measured different dispute patterns, e.g., disputing sent messages or disputing gates of the `computeRound` functions, and picked the pattern with the highest costs. The execution costs, both optimistic and worst case, incorporate all protocol steps, incl. the secure

funding of the instance. We exclude the derivation of the initial seeds as this step strongly depends on the underlying PVC protocol.

In the optimistic case, the costs of executing our protocol are similar to the ones of [ZDH19]. The authors report a gas consumption of 482k gas while our protocol consumes between 465k and 1M gas, depending on the number of parties – recall that the protocol of [ZDH19] is restricted to the two-party setting. This overhead in our protocol when considering more than two parties is mainly introduced by the fact that [ZDH19] does assume a single deposit while our implementation requires each party to perform a deposit.

Unfortunately, we cannot compare worst-case costs directly, as the protocol of [ZDH19] validates the consistency of a fixed data structure, i.e., a garbled circuit, while our implementation validates the correctness of the whole protocol execution. In particular, [ZDH19] performs a bisection over the garbled circuit while we perform two bisections, first over the message history and then over the computation generating the outgoing messages; such a message might for example be a garbled circuit. Further, [ZDH19] focuses on a boolean circuit, while we model the `computeRound` function as an arithmetic circuit – as the EVM always stores data in 32-byte words, it does not make sense to model the function as a boolean circuit. Although not directly comparable, we believe the protocol of [ZDH19] to be more efficient for the special case of a two-party garbling protocol, as the protocol can exploit the fact that a dispute is restricted to a single message, i.e., the garbled circuit, and the data structure of this message is fixed such that the dispute resolution can be optimized to said data structure.

Our measurements indicate that the worst-case costs of each scenario are always defined by a dispute pattern that does not dispute a message chunk but a gate of the `computeRound` functions. This is why the message chunks have no influence on the worst-case execution costs. Of course, this observation might be violated if we set the number of chunks much higher than the number of gates. However, it does not make sense to have more message chunks than gates because each message chunk needs to be mapped to a gate of the `computeRound` function defining the value of said chunk.

Both, the number of rounds and the number of parties increase the maximal size of the disputed message history and, hence, the depth of the bisected history tree. As the depth of the bisected tree grows logarithmic in the tree size, our protocol is highly scalable in the number of parties and rounds.

Finally, we note that we understand our implementation as a research prototype showing the practicability of our protocol. We are confident that additional engineering effort can further reduce the gas consumption of our contract.

Acknowledgments. The first, third, and fourth authors were supported by the German Federal Ministry of Education and Research (BMBF) *iBlockchain project* (grant nr. 16KIS0902), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) *SFB 1119 - 236615297 (CROSSING Project S7)*, by the BMBF and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the *National Research Center for Applied Cybersecurity ATHENE*, and by Robert Bosch GmbH, by the Economy of Things Project. The second author was supported by the BIU Center for Research in Applied Cryptography and Cyber

Security in conjunction with the Israel National Cyber Bureau in the Prime Minister's Office, and by ISF grant No. 1316/18.

References

- [ADMM14] Andrychowicz, M., Dziembowski, S., Malinowski, D., Mazurek, L.: Secure multiparty computations on bitcoin. In: IEEE SP (2014)
- [AL07] Aumann, Y., Lindell, Y.: Security against covert adversaries: efficient protocols for realistic adversaries. In: TCC (2007)
- [AO12] Asharov, G., Orlandi, C.: Calling out cheaters: covert security with public verifiability. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 681–698. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34961-4_41
- [BK14] Bentov, I., Kumaresan, R.: How to use bitcoin to design fair protocols. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8617, pp. 421–439. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44381-1_24
- [BMR90] Beaver, D., Micali, S., Rogaway, P.: The round complexity of secure protocols (extended abstract). In: STOC (1990)
- [Coi21] CoinMarketCap. Ethereum (ETH) price (2021). <https://coinmarketcap.com/currencies/ethereum/>
- [CRR11] Canetti, R., Riva, B., Rothblum, G.N.: Practical delegation of computation using multiple servers. In: CCS (2011)
- [DKL+13] Damgård, I., Keller, M., Larraia, E., Pastro, V., Scholl, P., Smart, N.P.: Practical covertly secure MPC for dishonest majority – or: breaking the SPDZ limits. In: Crampton, J., Jajodia, S., Mayes, K. (eds.) ESORICS 2013. LNCS, vol. 8134, pp. 1–18. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40203-6_1
- [DOS20] Damgård, I., Orlandi, C., Simkin, M.: Black-box transformations from passive to covert security with public verifiability. In: Micciancio, D., Ristenpart, T. (eds.) CRYPTO 2020. LNCS, vol. 12171, pp. 647–676. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-56880-1_23
- [DPSZ12] Damgård, I., Pastro, V., Smart, N., Zakarias, S.: Multiparty computation from somewhat homomorphic encryption. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 643–662. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32009-5_38
- [EFS20] Eckey, L., Faust, S., Schlosser, B.: OptiSwap: fast optimistic fair exchange. In: ASIA CCS (2020)
- [Eth21] Etherscan. Ethereum Average Gas Price Chart (2021). <https://etherscan.io/chart/gasprice>
- [FHKS21] Faust, S., Hazay, C., Kretzler, D., Schlosser, B.: Generic compiler for publicly verifiable covert multi-party computation. In: Canteaut, A., Standaert, F.-X. (eds.) EUROCRYPT 2021. LNCS, vol. 12697, pp. 782–811. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77886-6_27
- [GMW87] Goldreich, O., Micali, S., Wigderson, A.: How to play ANY mental game or a completeness theorem for protocols with honest majority. In: STOC (1987)

- [HKK+19] Hong, C., Katz, J., Kolesnikov, V., Lu, W., Wang, X.: Covert security with public verifiability: faster, leaner, and simpler. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019. LNCS, vol. 11478, pp. 97–121. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17659-4_4
- [KB14] Kumaresan, R., Bentov, I.: How to use bitcoin to incentivize correct computations. In: CCS (2014)
- [KGC+18] Kalodner, H.A., Goldfeder, S., Chen, X., Matthew Weinberg, S., Felten, E.W.: Arbitrum: scalable, private smart contracts. In: USENIX Security (2018)
- [KM15] Kolesnikov, V., Malozemoff, A.J.: Public verifiability in the covert model (almost) for free. In: Iwata, T., Cheon, J.H. (eds.) ASIACRYPT 2015. LNCS, vol. 9453, pp. 210–235. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-48800-3_9
- [SSS21] Scholl, P., Simkin, M., Siniscalchi, L.: Multiparty computation with covert security and public verifiability. IACR Cryptology ePrint Archive (2021)
- [TR19] Teutsch, J., Reitwießner, C.: A scalable verification solution for blockchains. CoRR, abs/1908.04756 (2019)
- [W+14] Wood, G., et al.: Ethereum: A secure decentralised generalised transaction ledger. Ethereum project yellow paper (2014)
- [WRK17a] Wang, X., Ranellucci, S., Katz, J.: Authenticated garbling and efficient maliciously secure two-party computation. In: CCS (2017)
- [WRK17b] Wang, X., Ranellucci, S., Katz, J.: Global-scale secure multiparty computation. In: CCS (2017)
- [YWZ20] Yang, K., Wang, X., Zhang, J.: More efficient MPC from improved triple generation and authenticated garbling. In: CCS (2020)
- [ZDH19] Zhu, R., Ding, C., Huang, Y.: Efficient publicly verifiable 2pc over a blockchain with applications to financially-secure computations. In: CCS (2019)