Paralysis Proofs: Secure Dynamic Access Structures for Cryptocurrencies and More

Fan Zhang  
Cornell University

Philip Daian  
Cornell Tech

Gabriel Kaptchuk  
Johns Hopkins University

Iddo Bentov  
Cornell University

Ian Miers  
Cornell Tech

Ari Juels  
Cornell Tech

Abstract

Conventional \((M, N)\)-threshold signature schemes leave users with a painful choice. Setting \(M = N\) offers maximum resistance to key compromise. With this choice, though, loss of a single key renders the signing capability unavailable, creating paralysis in systems that use signatures for access control. Lower \(M\) improves availability, but at the expense of security. For example, a \((3, 3)\)-multisignature cryptocurrency wallet experiences access-control paralysis upon loss of a single key, but a \((2, 3)\)-multisig allows any two players to collude and steal funds from the third.

In this paper, we introduce techniques that address this impasse by making general cryptographic access structures dynamic. Our schemes permit, e.g., a \((3, 3)\)-multisig, to be downgraded to a \((2, 3)\)-multisig if a player goes missing. This downgrading is secure in the sense that it occurs only if a player is provably unavailable.

Our main tool is what we call a Paralysis Proof, evidence that players, i.e., key holders, are unavailable or incapacitated. Using Paralysis Proofs, we show how to construct a Dynamic Access Structure System, which can securely and flexibly update target access structures without a trusted third party such as a system administrator. We present DASS constructions that combine a trust anchor (a trusted execution environment or smart contract) with a censorship-resistant channel in the form of a blockchain. We offer a formal framework for specifying DASS policies, and define and show how to achieve critical security and usability properties (safety, liveness, and paralysis-freeness) in a DASS.

Paralysis Proofs can help address pervasive key-management challenges in many different settings. We present DASS schemes for three important example use cases: recovery of cryptocurrency funds should players become unavailable, returning funds to users when cryptocurrency custodians fail, and remediating critical smart-contract failures such as frozen funds. We report on practical implementations for Bitcoin and Ethereum.

1 Introduction

A common paradox in key management and access control systems is the “always/never” dilemma: systems must be always available when authorized, and never used when not [71]. In general, this trade-off between availability and potential for misuse exists in all key management or authentication systems, extending even to nuclear arms [28]. One simple example is the task of securing a single private key: replicate the key broadly across geography, machine architectures, and custodians, and the attack surface for key compromise increases. Store the key in a single secured location and the probability of loss increases, as do barriers to timely access.

Generally, this challenging trade-off is navigated through organizational structures for distributing control to a small number of trusted and independent parties, thereby choosing the desired optimum point on the defensibility vs. availability spectrum. Sometimes, however, this is impossible. In a cryptocurrency wallet setting, for example, large direct incentives for theft combined with suspicion among users make distributing trust extremely difficult. Worse still, a number of access-control failures have occurred through both custodial failures [11, 62] and user errors [22, 64, 81].

Consider for example three players, Alice, Bob, and Carol, who jointly own a pool of cryptocurrency. For maximum protection against key compromise, they might use a \((3, 3)\)-multisignature wallet ("multisig"). Spending the money would then require the signatures of all keys \(s_{k_A}, s_{k_B}\) and \(s_{k_C}\). This provides good security: an adversary would need to compromise all three keys to steal the money. If, however, just one key became unavailable, e.g., Alice lost \(s_{k_A}\), the money would be permanently lost. To provide availability should a key be lost, the players might instead use a \((2, 3)\)-multisig. Unfortunately, in this case, any two players could cheat the third.

Neither choice seems optimal: it appears that these three players can realize either good security or good availability.

In this paper, we show how to achieve both properties and obtain a better security vs. availability trade-off than either multisig option alone. The key idea is to enable an multisig to be conditionally downgraded, e.g., to change a \((3, 3)\)-multisig to a \((2, 3)\)-multisig only if a player becomes unavailable. More generally, we show how to securely downgrade (or otherwise change) any access structure [37], i.e., policy determining which players can control a resource. Our system does not completely avoid security/availability trade-offs, as some lengthy interactive challenges may be required to access funds. It does however enable new and desirable points on the trade-off spectrum that provably prevent against permanent availability loss while providing users maximum security, very desirable properties in high-assurance control mechanisms for e.g. cryptocurrencies.

The main technical challenge in our work is ensuring that downgrading only happens when a player is truly unavailable. Otherwise, players can cheat by simulating the disappearance of a live player.
Secure downgrading relies crucially on an ability to construct strong proofs of player incapacitation, what we call Paralysis Proofs.

Constructing and consuming Paralysis Proofs securely without a trusted third party (TTP) such as a system administrator, as we aim to do in this paper, is challenging, as explained below. We show how two new technologies, trusted execution environments (TEEs) and blockchains, together make it feasible for the first time.

**Problem setting** Paralysis Proofs help address a fundamental key-management challenge that is pervasive in cryptographic systems. Cryptocurrencies vividly illustrate the problem.

Private signing keys for cryptocurrencies embody direct and total control over funds. Key theft thus leads to immediately and irrevocably stolen money, and is therefore attractive for hackers. Over 980,000 Bitcoin (worth about $9.8 billion at the time of writing) [68] have been stolen from exchanges alone.

At the same time, lost keys—resulting from, e.g., accidental deletion or corruption—can also be catastrophic. It is estimated that nearly 4 million Bitcoin (worth around $4 billion at the time of writing), have been lost forever due to lost keys [69]. Unlike traditional online banking systems where lost credentials can be recovered via out-of-band mechanisms, the decentralized nature of cryptocurrencies makes direct recovery of credentials from these systems impossible.

The main approach to dealing with such problems in cryptocurrencies is multisig addresses or wallets [60]; more generally, secret sharing is a common technique [73]. These techniques enable general access structures in which a set of players holds key shares, and predetermined subsets of players can use their shares to access the target resource (e.g., spend cryptocurrency). The most common choice is a simple (M, N)-access structure (for M ≤ N), which gives control to any subset of M players among a total set of N. For a given N, M dictates a trade-off between security and availability. Larger M means higher security, as more players must agree to access the resource, while smaller M means fewer players must agree, lowering per-response effort and implying higher availability.

Today’s systems, however, only support static or fixed access structures (i.e., for all times t1, t2, any set of players that satisfies the access structure at time t1 also does so at time t2). When no choice of M offers a good security vs. availability balance, as in our example, users face an unsolvable dilemma.

Paralysis Proofs help resolve this dilemma. They enable secure dynamic access structures in what we call a Dynamic Access Structure System (DASS), a system that changes M securely as players become unavailable. We call this process access-structure migration, as access may be changed beyond simple downgrades.

**Overview of our Approach** Proving to a third party that a player is available is easy: Just have her sign a fresh message. But how can Bob prove that Alice cannot (or refuses to) sign a message, as opposed to Bob faking a failure and not communicating with Alice?

Our Paralysis Proof constructions leverage the censorship-resistance [27] and data persistence of a public blockchain to detect and record the fact of an unavailable player. If, e.g., Alice disappears, a challenge can be issued to her on chain. Public blockchains are censorship-resistant, in the sense that if Alice tries to post a response within Δ epochs (blocks) for some suitable Δ she can do so with high probability even in the face of powerful network adversaries. Thanks to this property, lack of response from Alice within Δ blocks of a challenge constitutes a Paralysis Proof showing Alice’s unavailability.

Given a TTP, it is then easy to migrate an access structure securely. For example, our three players might use a (3, 3)-multisig, with skA, skB and skC held by the TTP. Given a Paralysis Proof showing Alice has disappeared, the TTP could release skA to Bob or Carol, effectively downgrading to a (2, 3)-multisig.

We aim in our work, however, to avoid TTPs. Our DASS schemes therefore rely on two technologies that serve as trust anchors by emulating TTPs: smart contracts where available, and trusted execution environments (TEEs) such as Intel SGX, where they are not.

**Example applications** We explore three illustrative use cases for Paralysis Proofs, exploring forms of paralysis:

- **Cryptocurrency key loss:** Generalizing our running example, we present protocols that permit any (M, N)-multisig cryptocurrency scheme to be downgraded (i.e., have lower M) if and only if keys or players go missing. We report on a purely smart-contract based Ethereum implementation and one for Bitcoin that involves use of a TEE. The Bitcoin scripting language, despite its time-based opcodes, cannot support a Paralysis Proof System; we present some script-based alternative schemes, however, that are less secure than our main scheme and/or require proposed enhancements to Bitcoin.

- **Cryptocurrency custody failures:** The challenges of key maintenance and endpoint security make third-party custody solutions appealing to users. Custodians can themselves become paralyzed, though, and can also freeze specific users’ funds. We propose a DASS which, if a custodian is failing to honor a valid withdrawal request for any reason, allows a user to migrate control of her funds to a backup key. Because this key is not usable for spending absent such a proof, the user is not trusting her own endpoint’s key management security by default.

- **Smart contract failures:** Smart contract bugs can result not just in exploits, but in funds being permanently paralyzed. The famous (second) Parity Multisig Wallet bug is just one example that permanently froze approximately $150 million in Ether in 2017 [65]. We propose a continuous-integration framework that regularly applies a test suite to a smart contract to validate its correct functioning, including liveness of funds. If (and only if) fund paralysis occurs, a paralysis proof submission triggers an "escape hatch" [51], failover logic that refunds or moves a smart contract’s assets. Our techniques are general, and can help address other non-paralysis smart contract failures.

We emphasize that Paralysis Proofs are general and can be applied beyond cryptocurrency, e.g. for (M, N)-sharing of decryption keys.

We also emphasize that a TEE-based application can, of course, store a master key or all players’ keys and directly mediate all access requests by players. Our schemes, however, avoid placing...
a TEE on the critical path for ordinary transactions, preventing vulnerability to denial-of-service and service or hardware failures.

Contributions

In summary, our main contributions are as follows:

- **Paralysis Proofs**: We introduce the concept of Paralysis Proofs, and show how to achieve them using blockchains. We also introduce Dynamic Access Structure Systems (DASSes), which combine Paralysis Proofs with a trust anchor (smart contract or TEE) to enable secure access-structure migration. They thereby offer security vs. availability trade-offs unachievable in conventional, static access-structure systems.

- **Formal definitions and framework**: We formally define key properties for a DASS (liveness, safety) and its underlying migration policy, called a DASP (privilege-preserving, paralysis-free) (Section 3). We also present a general Universal-Composability-type ideal functionality that formally specifies security properties required for a broad range of applications (Section 4).

- **Applications**: We present three example applications: Cryptocurrency key loss (Section 4), cryptocurrency custody failures (Section 5), and smart contract failures (Section 6).

- **Implementation**: We present implementations in Ethereum and Bitcoin for the first application, using smart contracts and TEEs (Intel SGX, in particular) respectively. We offer a UC proof (sketch) of security for the latter. To illustrate the limitations of pure blockchain approaches for Bitcoin, we also explore script-based schemes in the paper appendix.

- **TEE compromise**: We explore alternative DASS designs that provide resilience to TEE compromise, such as through side-channel attacks demonstrated against SGX (Section 4.5).

2 Background

In this section we provide some basic background on Trusted Execution Environments, Bitcoin, and smart contracts.

**Trusted Execution Environments and SGX** A Trusted Execution Environment (TEE) is a black-box-like execution environment that provides confidentiality and integrity for applications running on potentially malicious hosts.

Intel Software Guard Extension (SGX) [6, 34, 36, 54] is a realization of TEEs as a new instruction set architecture extension enabled on most new Intel processors. SGX allows processes to execute in an *enclave*, an environment that enforces application confidentiality and integrity against even a malicious operating system and some classes of hardware attacks. SGX also enables applications to emit third-party verifiable *attestations* to their origin and outputs. Enclaves cannot make system calls, but can communicate with untrusted programs running in the host OS. As a result, an enclave depends upon a potentially malicious operating system for network and file system operations. Enclaves can therefore secure applications’ state and execution, but cannot ensure successful network communications or file accesses. Despite these limitations, enclaves are powerful tools for building a variety of functionalities [21, 40, 91].

**Bitcoin** Bitcoin is a decentralized electronic cash scheme in which transactions moving funds are recorded in an append only log, a blockchain. Rather than storing funds in accounts whose balance is altered by transactions, Bitcoin uses transactions themselves to record both ownership and balance. Transactions consist of *inputs* and *outputs*. An output consists of an amount and a *script_pubkey* that specifies how that amount can be spent. Inputs specify the transaction output which is the source of the funds and include a *script_sig* showing authorization to use the funds. Thus transactions spend the outputs of previous transactions. Unconsumed outputs are known as *UTXOs* or Unspent Transaction Outputs. One can check if a transaction has been spent by seeing if it is in the set of UTXOs. By requiring that outputs can only be spent once and that the amount of money included in a transaction’s inputs is at least as much as its outputs, Bitcoin enforces invariants of a monetary system and prevents forgery.

Beyond monetary invariants, Bitcoin must also handle access control. *script_pubkey* and *script_sig* are the authorization mechanisms for transactions that ensure funds cannot be stolen or misused. Typically the *script_pubkey* in an output specifies the keys that must sign any transaction spending that output. This may be a single key or an arbitrary combination of keys, e.g., $(pk_1 \land pk_2) \lor pk_3$. An input consuming an output with such a *script_pubkey* would then need a signature that satisfied that requirement, e.g., it would need to contain signatures under both $pk_1$ and $pk_2$. These requirements are represented in a stack based language known as *Bitcoin script*. While in principle Bitcoin script can represent complex logic, in practice limitations on supported instructions and the length of a script mean it is mainly used for simple authorization checks.

**Smart Contracts** Smart contracts are small, deterministic programs that are stored in a blockchain system’s state and interpreted by a virtual machine. Beyond the value field associated with a simple currency transaction, transactions on smart contract blockchains contain two additional key parameters: *input data* and *code location* [87]. To process a transaction, a smart contract system looks up the code stored in the provided location, executing the code with the provided input data and processing any side effects output by the execution of the code. A key differentiator between such smart contract-enabled systems and simpler script like Bitcoin script is the capability of contract platforms to provide rich-statefulness [85]. In a system providing rich statefulness, all executing transactions have native access to persistent state (stored across transactions, blocks, and time), and can interact with and update both local state for data storage and global blockchain state for system-wide information (like the current height of the blockchain, or the hash of the block the transaction eventually gets mined in). Rich statefulness is particularly relevant to our system, as access to state across time is required to track whether a Paralysis Proof has been initiated, is pending, or has been responded to. It is Bitcoin’s lack of such native statefulness that makes trusted hardware the only practical solution for handling paralysis with large numbers of users.
3 Dynamic Access Structure Systems

In this section, we develop both formal definitions and a framework for analyzing security of Dynamic Access Structure Systems.

A Dynamic Access Structure Policy (DASP) consists of a set of access structures and rules dictating migration conditions among them. For example, "this Bitcoin fund requires signatures from Alice, Bob and Carol to spend; if any of them disappears, signatures from the remaining two suffice to spend the fund" informally specifies a DASP. The access structure is the set of holders authorized to spend Bitcoin, and migration entails removing unresponsive signers.

We use the term Dynamic Access Structure System (DASS) to denote a system that enforces a DASP. Essential to our DASS constructions is the use of Paralysis Proofs to demonstrate conditions, e.g., party incapacitation, that justify migration from one access structure to another. We now provide formalism for DASP specification (Section 3.1), followed by security definitions for a DASS (Section 3.2).

3.1 Specification of a Dynamic Access Structure Policy

3.1.1 Basic Definitions A Dynamic Access Structure Policy (DASP) comprises a tuple \( (R, S, M) \) that specifies the resources \( R \) being access-controlled, a set of access structures \( S \), and a set of migration rules \( M \) dictating conditions under which access-structure migrations are permitted.

Let \( \{P_i\}_{i=1}^N \) denote the set of \( N \) parties at beginning of the protocol, and \( L_t \) the set of live (i.e., not incapacitated) parties at time \( t \). As we shall see shortly, correctly determining \( L_t \), i.e. which parties are actually live, is the main technical challenge in enforcing a DASP. We use \( L_t \) to denote the ground truth. We assume that if a party becomes incapacitated, it remains incapacitated throughout the protocol, i.e. \( P \notin L_t \) implies \( P \notin L_{t'} \) for all \( t' > t \).

In this paper, an access structure \( s \) is a function \( s(L) \rightarrow \{\text{true}, \text{false}\} \) that determines whether a set of live parties \( L \subseteq \{P_i\} \) is allowed to access the managed resource. Access structures are monotonic, i.e., \( s(L) = \text{true} \) and \( L \subseteq L' \) together imply that \( s(L') = \text{true} \). A migration rule \( m_{s_i, s_j} \in M \) is a function \( m_{s_i, s_j}(L) \rightarrow \{\text{true}, \text{false}\} \) that determines whether migrating from \( s_i \) to \( s_j \) is permitted if the set of live parties is \( L \). We use \( L \xrightarrow{L} s_j \) to denote \( m_{s_i, s_j}(L) = \text{true} \).

For a given DASP, the set of access structures \( S \) and the associated migration rules \( M \) may be represented as a directed graph \( G = (S, M) \). Here we overload \( S \) and \( M \) to denote respectively the sets of nodes and edges. A node \( s_i \in S \) is an access structure and an enhanced edge \( (s_i, s_j) \in M \) represents the migration rule \( m_{s_i, s_j} \), which specifies the condition to migrate from access structure \( s_i \) to \( s_j \). Access structure \( s_0 \) is reachable from \( s_j \) by \( L \subseteq L_0 \), denoted \( L \xrightarrow{s_j} s_0 \), if there exists a path \( (s_1, s_2, \ldots, s_n) \) in \( G \) such that for all \( i \in [1, n-1] \), \( m_{s_i, s_{i+1}}(L) = \text{true} \).

3.1.2 Security Goals. A fundamental correctness requirement for any access control is that migration between access structures does not eliminate the privilege of live parties. We capture this notion by stipulating that a DASP be privilege-preserving. To define this property, we first require two technical definitions.

**Definition 3.1.** The set of least permissive access structures for \( L \subseteq \{P_i\} \), denoted by \( S_{LP}(L) \), is as follows:

\[
S_{LP}(L) = \{s \in S : s(L) = \text{true} \land (\forall L' \subseteq L, s(L') = \text{false})\}.
\]

Intuitively, \( S_{LP}(L) \) is the set of all access structures such that if the only possible live parties are in \( L \), then all such parties must be live to access the resource. Given this, we define privilege-preserving:

**Definition 3.2. (Privilege-preserving)** Let \( L_1 \) be the set of live parties at time \( t \). A DASP \( (R, S, M) \) is privilege-preserving if \( L_1 \) can never migrate to an access structure that can be satisfied with a set \( L' \) of parties such that \( L \nsubseteq L' \) at any time \( t \). Formally, \( \forall s \in S \) such that \( s(L_1) = \text{true} \):

\[
\forall s' \text{ s.t. } s \xrightarrow{L_1} s', s' \in \bigcup L_1 \subseteq S_{LP}(L).
\]

A DASP is paralysis-free if, when the current access structure cannot be satisfied, switching to another satisfiable access structure is permitted provided that the migration will not deprive the privilege of any live party.

**Definition 3.3. (Paralysis-freeness)** Let \( L_1 \) be the set of live parties at time \( t \). A DASP \( (R, S, M) \) is paralysis-free if at any time \( t \), \( \forall s \in S \) such that \( s(L_1) = \text{false} \):

\[
S_{LP}(L_1) \neq \emptyset \quad \exists s' \in S_{LP}(L_1) \text{ s.t. } s \xrightarrow{L_1} s'.
\]

Note that a paralysis-free DASP doesn’t imply the availability of the resource. What a paralysis-free policy can guarantee is the best possible availability: if there is a access structure that can get the system out of paralysis, then the DASP should permit a transition to that access structure. However, if the set of live parties is too sparse to satisfy any of the prescribed access structures, then the desired availability cannot be achieved.

**Example 3.4.** Let’s take the example of \( N \) shareholders who wish to retain access to the resource \( R \) should one party disappear. Let \( \mathcal{P} = \{P_i\}_{i=1}^N \) denote the set of \( N \) parties, and \( \mathcal{P}_{-i} = \mathcal{P} \setminus \{P_i\} \) denote the set of \( N-1 \) parties that excludes \( P_i \). Let \( \mathcal{I}(-) \) denote an indicator function. A DASP \( (R, S, M) \) that realizes the aforementioned access control can be specified by \( S = \{s_i\}_{i=0}^N \) where

\[
s_0 = \mathcal{I}_i \quad s_i = \mathcal{I}_{P_{-i}}, 1 \leq i \leq N
\]

and the condition \( m_{s_0, s_i} \in M \) is fulfilled for \( L_1 = \mathcal{P}_{-i} \).

According to Definition 3.2 and Definition 3.3, the DAS in Example 3.4 is privilege-preserving and paralysis-free.

3.2 Security definitions for a DASS

We use the term Dynamic Access Structure System (DASS) to denote a system that enforces a DASP. In this section, we formally define the security of a DASS with a Universal Composability (UC) [19] ideal functionality \( F_{DASS} \). Later in Section 4 we present a protocol \( \Pi_{SGX} \) that UC-realizes \( F_{DASS} \).
3.2.1 Adversarial model. We assume an adversary that may corrupt an arbitrary number of parties. An honest party always follows the protocol, while a corrupted party controlled by the adversary may deviate arbitrarily (i.e., Byzantine corruption). We assume that the adversary has complete control of the network, with the exception that a blockchain is available to all parties, i.e., is censorship-resistant, and the maximum network latency to the blockchain is bounded by a known $\Delta$.

3.2.2 Ideal Functionality. We specify security goals of a DASS in the ideal functionality $\mathcal{F}_{\text{DASS}}$ as defined in Figure 1.

To reduce clutter, we omit the handling of session IDs [19] in $\mathcal{F}_{\text{DASS}}$ but readers are advised that messages received and sent by $\mathcal{F}_{\text{DASS}}$ are implicitly associated with an SID. When $\mathcal{F}_{\text{DASS}}$ sends subroutine output to parties, we use the delayed output terminology from [19] to model the power of the network adversary. Specifically, when $\mathcal{F}_{\text{DASS}}$ sends a public delayed output to party $P_i$, the output is first sent to $\mathcal{A}$ and then forwarded to $P_i$ after $\mathcal{A}$'s acknowledgement or $\Delta$ time has elapsed, whichever happens first.

$\mathcal{F}_{\text{DASS}}$ maintains internal states $(L_t, s)$ for the set of live parties and currently enforced access structure respectively. To capture the paralysis explicitly, we extend the standard corruption model [19] with special "paralysis" corruption. Upon receipt of a paralysis message from $\mathcal{A}$, a party immediately announces its paralysis and halts until the end of the protocol. In the ideal protocol, $\mathcal{A}$ sends (paralysis, $P_i$) to $\mathcal{F}_{\text{DASS}}$, who removes $P_i$ from the set of live parties.

To access the resource, a set of parties $P$ send (access, inp), in which inp specifies the parameter of access, to $\mathcal{F}_{\text{DASS}}$. If $P$ is permitted to access by the current access structure, i.e. $s(P) = \text{true}$, $\mathcal{F}_{\text{DASS}}$ returns the result of accessing $R$. A set of parties can initiate a migration to another access structure $s'$ by sending (migrate, $s'$) to $\mathcal{F}_{\text{DASS}}$. If the transition to $s'$ is permitted by the enforced DASP, $\mathcal{F}_{\text{DASS}}$ sets the current enforced access structure to $s'$.

3.2.3 Security Properties $\mathcal{F}_{\text{DASS}}[s_0, \mathcal{R}, s, M]$ encapsulates the following security properties of a Dynamic Access Structure System. Let $s$ denote the effective access structure of $\mathcal{F}_{\text{DASS}}$, and $L_t$ the set of live parties at time $t$. $\mathcal{F}_{\text{DASS}}$ guarantees both safety and liveness in all states $s \in S$ at any time $t$:

**Safety:**
- A set of parties $L \subseteq L_t$ can access $\mathcal{R}$ only if $s(L) = \text{true}$.
- A transition to $s' \neq s$ occurs only if $m_{s, s'}(L_t) = \text{true}$.

**Liveness:**
- If $s(L) = \text{true}$ for some $L \subseteq L_t$, then $L$ can access $\mathcal{R}$ within $\Delta$ time after interacting with the DASS honestly.
- If $m_{s, s'}(L_t) = \text{true}$, then a transition to $s' \neq s$ occurs within $\Delta$ after $L_t$ interacts with the DASS honestly.

**Examples** Consider a DASS enforcing the DASP in Example 3.4, the Safety property ensures that access is enforced by the current access structure at any time, and that the access structure can be downgraded to allow access by $N-1$ shareholders only if $|L_t| < N$.

The Ideal Functionality of a Dynamic Access Structure System $\mathcal{F}_{\text{DASS}}[s_0, \mathcal{R}, s, M]$ with parties $(P_i)_{i \in [n]}$

1. On receiving (init) from any $P_i$:
2. If $|P_i| = |L_0|$, sets $s := s_0$
3. On receiving (paralysis, $P_i$) from $\mathcal{A}$:
4. $L_t = L_t \setminus \{P_i\}$
5. On receiving (access, inp) from $P_j \in L_t$:
6. let current time be $t$
7. if no access request for inp or it has expired, create a new one
8. else: store (inp, $(P_i, t)$), overwriting (inp, s, c) if exists
9. if $s(P) = \text{true}$ then:
10. send a public delayed output $R(P, \text{inp})$ to all parties in $P$
11. On receiving (migrate, $s'$) from $P_j \in L_t$:
12. $s' \in S \land m_{s, s'} \in M$
13. $L_{\text{fake-death}} = \emptyset$
14. for all corrupted parties $P_e \in L_t$:
15. ask $\mathcal{A}$ if $P_e$ choose to pretend to be paralyzed
16. if so add $P_e$ to $L_{\text{fake-death}}$
17. if $m_{s, s'}(L_t \setminus L_{\text{fake-death}}) = \text{true}$ then:
18. sends a public delayed output $(s', \text{ok})$ to all parties; $s = s'$

Figure 1: The ideal functionality of a Dynamic Access Structure System. The entry point marked with * is only executed once.

4 Paralysis Proofs for Cryptocurrency

In this section, we expand on the use of Paralysis Proofs to recover from cryptocurrency key loss (and related failures, e.g., player disappearance). We focus on Bitcoin, which presents particular technical challenges. For comparison, we also briefly present a conceptually straightforward implementation for Ethereum.

i.e., a collusion of $N - 1$ shareholders cannot maliciously accuse the $N^{th}$ shareholder of being incapacitated and thereby steal her share. The Liveness property ensures that access is granted if the structure is satisfied by a set of cooperating parties. Moreover, if allowed by the policy, the Liveness property ensures that the access structure will be downgraded within a bounded time should parties submit legitimate requests. Note that the Liveness property does not stipulate that access structure $s_i$ for $i > 0$ is automatically instantiated if $|L_t| < N$. This is because parties may not immediately activate an access-structure migration; in fact, if all parties are incapacitated, such migration cannot happen.

3.2.4 DASSes and Paralysis Proofs. The main challenge in realizing $\mathcal{F}_{\text{DASS}}$ is to determine the set of live parties $L_t$ in a trustworthy way. Our solution to that is Paralysis Proofs. Specifically, a DASS realizing $\mathcal{F}_{\text{DASS}}$ leverages the censorship-resistance of blockchain to enable parties to construct Paralysis Proofs to prove that $P_i \notin L_t$ for a given party $P_i$ or similar facts about $L_t$. In particular, as we shall see, to prove that $P_i \notin L_t$, parties issue to $P_i$ a challenge on the blockchain. If $P_i$ does not respond within some time $\Delta$, the challenge together with evidence of this failure to respond constitute a Paralysis Proof that proves $P_i \notin L_t$.
It is challenging to implement secure Paralysis Proofs compatible with the current Bitcoin protocol because of the limited expressiveness of Bitcoin scripts. (We show in Appendix D that it is possible were a proposed feature called “covenants” available [58].) We therefore explore Paralysis Proofs for Bitcoin using TEEs—specifically, Intel SGX, a powerful TEE available in existing Intel CPUs. Readers can refer to Section 2 for background on SGX.

We give technical preliminaries and discuss our trust model in Section 4.1. We present our main protocol (denoted $\Pi_{SGX}$) in Section 4.2, and discuss its security in the Universal Composability framework in Section 4.4, giving a proof (sketch) in the paper appendix. We discuss ways to reduce trust assumptions for SGX nodes in Section 4.5. For comparison, we present our basic Dynamic Access Structure System for Ethereum in Section 4.6.

4.1 Preliminaries

Bitcoin Transaction and CSV. In Bitcoin, spendable money is known as an Unspent Transaction Output (UTXO). We use $(V, \phi)$ to denote an UTXO of $V$ coins and script $\phi$. The script $\phi$ stipulates the condition to be satisfied in order to spend the UTXO. A Bitcoin transition (with the exception of coinbase transitions) consumes a set of UTXOs and creates one or more new ones. We use \(\langle \text{In}_i \rangle_{i=1}^n \overset{w_1, \ldots, w_n}{\rightarrow} \langle \text{Out}_j \rangle_{j=1}^m\) to denote a Bitcoin transaction with $n$ inputs, $m$ outputs, and $n$ witnesses, one for each input, such that $w_i$ satisfies the script of $\text{In}_i$.

An essential ingredient of $\Pi_{SGX}$ is Bitcoin’s relative timeout script instruction, also known as CheckSequenceVerify or CSV [16]. By putting the CSV instruction with parameter $\tau$ in the script $\phi$ of a UTXO $u$, we assert that the transaction that spends $u$ must reside in a block whose height (or timestamp) is more than $\tau$ relative to $u$. This assertion can be part of a conditional script, such that other branches of the script do not need to satisfy the CSV condition.

SGX and Attested Execution. Throughout the paper, we use SGX as a concrete building block. However, our protocol can be realized by any TEE that protects confidentiality and integrity of computation, and can issue proofs (attestations) of computation correctness.

In our formal specification, we adopt the (local) ideal functionality $G_{SGX}$ by Pass et al [66] to model SGX. Informally, a party first loads a program $\text{prog}_{\text{encl}}$ into an SGX enclave with an install message. On a resume call, the program is run on the given input inp, generating an output along with an attestation $a_{SGX} = \Sigma_{SGX}.\text{Sig}(\text{skatt}.(\text{prog}_{\text{encl}}, \text{outp}))$, a signature under the hardware key skatt. The public key $pk_{\text{att}}$ is can be obtained from $G_{SGX}.\text{getpk}()$. We refer readers to [66] for details.

Ideal Blockchain. Our protocol relies on an append-only ledger. We define the ideal functionality $F_{\text{blockchain}}[\text{succ}]$ in Figure 6 (inspired by [20]) to model a general-purpose append-only ledger implemented by common blockchain protocols. The parameter $\text{succ}(\text{history}, \text{item}) \rightarrow \{0, 1\}$ is a function that specifies the criteria for a new item to be appended to history, modeling the notion of transaction validity. We retain the append-only property of blockchains but abstract away the inclusion of items in blocks.

![Figure 2: An SGX based protocol for Paralysis Proofs.](image)

We assume a trustworthy time source available to $F_{\text{blockchain}}$ and items are timestamped when added. In practice, block numbers will be accepted $\Delta$ time after $F_{\text{blockchain}}$ accept $t_2$.

Protocol $\Pi_{SGX}$ with $P_1, \ldots, P_N$

1: Hardcoded: $\delta$ (e.g. $10^{-4}$), network latency $\Delta$
2: For any party $P_i$:
3: On receiving $(\text{init})$ from environment $Z$
4: $(pk_i, sk_i) \leftarrow \text{KGen}(\text{in}^i)$; publish $pk_i$
5: wait to receive $\langle pk_j \rangle_{j \neq i}$ from other parties
6: send $(\text{init}, \text{pro}_{\text{encl}})$ to $G_{SGX}$ and wait to receive $\text{eid}$
7: send $(\text{eid}, \text{resume}, \text{init}, (pk_j)_{j=1}^N)$ to $G_{SGX}$, wait for $pk_{SGX}$; publish $pk_{SGX}$
8: send $(\text{init}, ((\delta, pk_{SGX}) \cdot V, \text{all } pk \in \{pk_j\}_{j=1}^N \lor pk_{SGX}))$ to $F_{\text{blockchain}}$
9: if $F_{\text{blockchain}}$ is not properly initialized: broadcast abort
10: else broadcast ok
11: wait to receive ok from others and abort if a abort is received
12: On receiving $(\text{access}, \text{addr}_{\text{new}})$ from environment $Z$
13: obtain $\text{UTXO}_{\text{hand}}$ from $F_{\text{blockchain}}$
14: compute $\sigma = \text{Sig}(sk_i, (\text{UTXO}_{\text{hand}}, \text{addr}_{\text{new}}))$
15: send $(\text{resume}, (\text{spend}, \sigma, \text{UTXO}_{\text{hand}}, \text{addr}_{\text{new}}))$ to $G_{SGX}$
16: On receiving $(\text{migrate}, P'_i)$ from environment $Z$
17: assert $P'_i \subseteq \{pk_i\}_{i=1}^N$
18: obtain $\text{UTXO}_{\text{hand}}$ from $F_{\text{blockchain}}$
19: send $(\text{resume}, (\text{migrate}, \text{UTXO}_{\text{hand}}, P'_i))$ to $G_{SGX}$ and wait for $t_1, t_2$
20: send $t_1, t_2$ to $F_{\text{blockchain}}$; $t_2$ will be accepted $\Delta$ time after $F_{\text{blockchain}}$ accept $t_1$.
Program for the SGX Enclave (prog\textsubscript{encl})

1: Hardcoded: $\delta$, $\epsilon$, network latency $\Delta$, access grace period $T_u$
2: On input (init, $P_0$):
3: 3 parties := $P_0$
4: ($sk\textsubscript{SGX}, pk\textsubscript{SGX}) \rightarrow KGen(1^s)$ and output $pk\textsubscript{SGX}$
5: On input (spend, $\sigma$, UTXO\textsubscript{fund}, addr\textsubscript{new}):
6: parse UTXO\textsubscript{fund} as $(V, (all pk \in P) \lor pk\textsubscript{SGX})$ or abort
7: if received $|P|$ requests for (UTXO\textsubscript{fund}, addr\textsubscript{new}) within $T_u$:
8:  assert $V(\sigma, pk_i)$ for all $1 \leq i \leq n$
9:  sign transaction $t := (UTXO\textsubscript{fund} \rightarrow addr\textsubscript{new})$ with $sk\textsubscript{SGX}$
10:  send $t$ to $\mathcal{F}_{blockchain}$
11: else store $\sigma$ and wait for more requests
12: On input (migrate, UTXO\textsubscript{fund}, $P'$):
13: parse UTXO\textsubscript{fund} as $(V, (all pk \in P) \lor pk\textsubscript{SGX})$ or abort
14: ($pk_k, sk_k) \rightarrow KGen(1^s)$
15: // a life signal for parties affected by the migration
16: $\phi_{\text{lifesignal}} := ((any pk \in P \lor V) \lor (pk_v \land (CSV \geq \Lambda)))$
17: sign transitions $t_1, t_2$ with $sk\textsubscript{SGX}$ and $pk_k$;
18: $t_1 := (\langle \delta, pk\textsubscript{SGX} \rangle \rightarrow (e, \phi_{\text{lifesignal}}), (\delta - \epsilon, pk\textsubscript{SGX}))$
19: $t_2 :=$
20: $(\langle e, \phi_{\text{lifesignal}}, V, (all pk \in P) \lor pk\textsubscript{SGX} \rangle \rightarrow (V, (all pk \in P) \lor pk\textsubscript{SGX}))$
21: output $t_1$ and $t_2$

Figure 3: The Paralysis Proof Enclave. The entry point marked with $*$ is only executed once.

**Initialization.** To start the protocol, some honest party needs to load an SGX instance with $prog\textsubscript{encl}$ and invoke the init procedure. For now we assume a single SGX available for all honest parties; thus any honest party can initiate the enclave (once initialized, sequential initializations will be ignored). In Section 4.5 we present an expanded distributed setup procedure that avoids the availability assumption and provides stronger guarantees.

After the setup procedure is completed, the parties send a small fund of $\delta \cdot B$ (e.g., $\delta = 0.00001$) to a new output that can be spent by $pk\textsubscript{SGX}$. Then the parties launch the protocol by sending their unspent output of $V$ coins (denoted UTXO\textsubscript{fund}) to a new output of $V$ coins with a script that can be spent by either $\{pk_i\}_{i=1}^N$ or $pk\textsubscript{SGX}$.

**Spending funds.** There are two ways to spend the funds that are managed in $\Pi\textsubscript{SGX}$. At any time, the players can spend the money via a Bitcoin transaction that embeds their $N$ signatures (per $\phi_{\text{life}}$ all in Figure 2). Hence, even in the case that all of the $N$ SGX CPUs are destroyed, the players are still able to spend the funds just as they could before the execution of $\Pi\textsubscript{SGX}$. However, a better way to spend the funds is by sending $\Delta \cdot B$ requests to an enclave, letting the enclave create a Bitcoin transaction with a single signature (signed by $sk\textsubscript{SGX}$). This reduces the on-chain complexity and the transaction fee (see also Appendix E).

**Migrating to another access structure.** The migrate procedure of $\Pi\textsubscript{SGX}$ resolves system paralysis by letting the live shareholders spend the money if one or more shareholders is incapacitated. Intuitively, the role of SGX is to be an arbitrator: when any shareholder alleges that the money is stuck due to an unresponsive party, SGX first gives the accused party $\Delta$ time to appeal, and the set of shareholders that controls the fund will be reduced only if no appeal is observed on the blockchain within this sufficiently large $\Delta$ (meaning that such an appeal did not occur, assuming censorship resistance [27] holds on the underlying blockchain).

The core idea of implementing an "appeal" in Bitcoin is to use what we call life signals. A life signal for party $P_k$ is a UTXO of negligible Bitcoin amount $\delta \cdot B$, that can be spent either by $P_k$—thereby signaling her liveness—or by $pk\textsubscript{SGX}$, but only after a delay. $\Pi\textsubscript{SGX}$ makes use of life signals to securely migrate to remove a party from the current access structure. Specifically, suppose the current set of shareholders is $P = \{P_i\}_{i=1}^N$, to (propose to) remove party $P_k$ from $P$, any live players can send a message (migrate, UTXO\textsubscript{fund}, $P \setminus \{P_k\}$) to $prog\textsubscript{encl}$. Then $prog\textsubscript{encl}$ will generate two signed transactions, $t_1$ and $t_2$ (defined in Figure 3 and illustrated in Figure 4), as follows:

- **Transaction $t_1$:** Acts as a life signal for $P_k$.
- **Transaction $t_2$:** Spends both the life signal (i.e., UTXO\textsubscript{lifeSignal}) and the escrowed fund (i.e., UTXO\textsubscript{fund}) to a script that is spendable without $P_k$ (i.e., by $\left(\{\{pk_{i}\}_{i=1}^N \setminus \{pk_k\}\}\right) \lor p_k\textsubscript{SGX}$).

The SGX enclave gives both $t_1$ and $t_2$ together as output. If $t_1$ is sent to the Bitcoin blockchain, $P_k$ can cancel her removal by spending $t_1$. Otherwise, $t_2$ will become valid after the $\Delta$ delay and can be sent to the blockchain, thereby removing $P_k$’s control over the fund. Figure 4 demonstrates an example with three players.

Notice that $prog\textsubscript{encl}$ parses UTxO\textsubscript{fund} and obtains the list of current shareholders, so that $prog\textsubscript{encl}$ does not have to keep track of current live shareholders locally, nor does it need to have an up-to-date view of the blockchain. As we’ll discuss shortly, this is an important security feature because it makes $\Pi\textsubscript{SGX}$ more resilient to SGX’s availability failure and avoids complexity of SGX with a blockchain.

4.3 Implementation

We implemented $\Pi\textsubscript{SGX}$ using Intel SGX SDK and Bitcoin Core. The source code is published at [2]. Our trusted functions contribute 874 lines of C++ code. The entire Trusted Computing Base (TCB) includes the Bitcoin Core implementation, two widely used cryptographic libraries (i.e., libsecp256k1 and OpenSSL), and the Panoply implementation described fully in [75].
4.4 Security of \( \Pi_{SGX} \)

Intuitively, the security of \( \Pi_{SGX} \) stems from the use of SGX and the relative timeout feature of Bitcoin. We first discuss the security of \( \Pi_{SGX} \) informally, then we present a formal security proof.

Use of relative timeout. The core security property of \( \Pi_{SGX} \) is that a live party cannot be falsely removed from the access structure, no matter how many of parties are malicious. This is achieved by the use of the relative timeout feature of Bitcoin [16] in the fresh \( t_1 \), and the atomicity of the signed transaction \( t_2 \).

To elaborate, \( t_2 \) will be valid only if the witness of both inputs (UTXO\textsf{fund} and UTXO\textsf{life signal}) is correct. The witness that the SGX enclave produced for spending the UTXO\textsf{fund} is immediately valid, but the witness for spending UTXO\textsf{life signal} becomes valid only after \( t_1 \) has been incorporated into a Bitcoin block that has been extended by \( \Delta \) additional blocks (due to the CSV condition). The shareholder \( P_i \) that accused \( P_k \) of being incapacitated should therefore broadcast \( t_1 \) to the Bitcoin network, wait until \( t_1 \) is added to the blockchain, then wait for the next \( \Delta \) blocks, and then broadcast \( t_2 \) to the Bitcoin network. However, while these \( \Delta \) blocks are being generated, \( P_k \) has the opportunity to appeal by spending \( t_1 \) with the secret key \( sk_k \) that is known only to her (the script of \( t_1 \) does not require the CSV condition for spending with \( sk_k \)). \( \Delta \) is set to a large enough value for two purposes: (1) to give \( P_k \) enough time to respond, and (2) to ensure that it is infeasible for an attacker to create a secretive chain of \( \Delta \) blocks faster than the Bitcoin miners, and then broadcast this chain (in which \( t_2 \) is valid) to overtake the public blockchain.

Note that a fresh, ephemeral key pair is generated for each life signal to ensure that \( t_1 \) is unique and hence does not already reside on the blockchain (e.g., \( P_k \) may have failed to respond to an earlier life signal but luckily another shareholder \( P_j \) was removed at that time). The SGX enclave does not need to store these ephemeral keys, as they are consumed right after generation.

In Appendix G we give a similar Paralysis Proof system that works with the current Bitcoin protocol and does not require SGX, but the construction has a weaker security guarantee and more than exponential overhead, substantially decreasing practicality.

No need to sync SGX with a blockchain. It is important to point out that the security of \( \Pi_{SGX} \) does not require the SGX enclave to have an up-to-date view of the blockchain (in fact it does not require any view of the blockchain), nor does it require a trusted clock. By contrast, protocols that require so have a larger attack surface, and in particular such protocols need additional security measures in order to be protected against rollback attacks (see, e.g., [10, 52]) and the problems of SGX clock (see, e.g. [20]).

Security Proof. The security of \( \Pi_{SGX} \) is formally analyzed using the framework developed in Section 3. Specifically, we formulate the security goal of \( \Pi_{SGX} \) as a DASP, and then prove in the University Composability (UC) framework that \( \Pi_{SGX} \) securely realizes the ideal functionality that implements the same DASP.

Specification of the DASP. The resource being managed by the \( \Pi_{SGX} \) is access to an oracle \( \mathcal{R}() \) that produces valid signatures authorizing Bitcoin expenditures. Initially, a Bitcoin fund is controlled by a set of \( N \) parties, denoted by \( P_0 \). \( \Pi_{SGX} \) aims to implement the following DASP specification:

\[
\begin{align*}
S & := \{ s_P() = I_P() : \forall P \in \mathcal{P}(P_0) \} \text{ and } \\
M & := \{ m_{sk, sk}() = \mathcal{I}_G() : A, B \subseteq P_0, B \subseteq A \} .
\end{align*}
\]

We prove that \( \Pi_{SGX} \) realizes (formally, UC-realizes) the ideal functionality \( \mathcal{F}_{DASP}[\mathcal{P}_0, \mathcal{R}, S, M] \) that enforces this DASP. In particular, we prove the following theorem:

**Theorem 4.1 (The Security of \( \Pi_{SGX} \)). Assume \( \mathcal{G}_{SGX} \)'s attestation scheme and the digital signature used in \( \Pi_{SGX} \) are existentially unforgeable under chosen message attacks (EU-CMA). Then \( \Pi_{SGX} \) UC-realizes \( \mathcal{F}_{DASP}[\mathcal{P}_0, \mathcal{R}, S, M] \) in the \((\mathcal{G}_{SGX}, \mathcal{F}_{blockchain})\)-hybrid model, for static adversaries.

**Proof.** See Appendix B for a proof sketch. □

4.5 Minimizing Trust in the TEE

We now briefly consider some ways to minimize the trust placed in the TEE (SGX node) employed in our protocol.

Avoiding a Single Point of Failure. Trusted hardware in general cannot ensure availability. In the case of SGX, a malicious host can terminate enclaves, and even an honest host could lose enclaves to outages. To avoid reliance on a centralized SGX server, each party in \( \Pi_{SGX} \) can run her own SGX enclave with an identical program. This way, any individual party (or set of parties) can always use all the capabilities of the protocol without being dependent on others.

Specifically, the initialization procedure of \( \Pi_{SGX} \) can be replaced with the following procedure that distributes the master key \( sk_{SGX} \) across multiple hosts. First, each enclave first generates a fresh key pair \( (pk_{SGX}, sk_{SGX}) \) and outputs \( pk_{SGX} \), while keeping \( sk_{SGX} \), secret. Then, each player uses her identity \( P_i \) to endorse \( pk_{SGX} \), and all the players reach agreement on the list of SGX identities \( \{ pk_{SGX}, i \} \). Finally, the enclaves then use \( \{ pk_{SGX}, i \} \) to establish secure channels (TLS) with each other, and create a fresh shared secret key \( sk_{SGX} \) that is associated with \( \{ pk_{SGX}, i \} \) (i.e., another invocation of the setup procedure will generate a different shared key).

Given use of the secure hardware random number generator (RDRAND), secret keys generated by SGX are known only to the enclaves, not to any of the players. From now on, no inter-enclave communication is needed in the course of the protocol. Each enclave then seals its state (which mainly consists of \( sk_{SGX} \)) by encrypting it using the hardware key (unique to each CPU) and storing the ciphertext to persistent storage. Hence, the enclave program does not have to run persistently, and each players can load and run the backup on-demand when needed.

Side-channel Resistance. Although SGX aims to provide confidentiality, recent work has uncovered data leakage via side-channel attacks [14, 31, 32, 39, 46, 47, 57, 63, 72, 86, 89]. Admittedly \( \Pi_{SGX} \) is not side-channel-free, but it has a relatively small and controlled attack surface. The only secret in SGX is \( sk_{SGX} \) and only operation involving \( sk_{SGX} \) is signature generation (besides key generation)—this makes \( \Pi_{SGX} \) amenable to software-level side-channel mitigations, such as constant-time ECDSA implementation (e.g. [24]).
A more powerful and somewhat more interesting approach is to design side-channel-free Paralysis Proofs. We claim that no side-channel-free construction of Paralysis Proofs exists given the current trust assumptions. However, if we relax the assumptions slightly, for example, by assuming a trusted relative clock in SGX, or assuming certain stationarity properties of the blockchain (e.g., difficulty), a side-channel-free Paralysis Proofs can be constructed by establishing an up-to-date view of blockchain in SGX (e.g., using techniques in [20]). Specifically, SGX will only be activated when paralysis happens (which requires an up-to-date view of the blockchain to detect), and will generate a new key $s_{SGX}$ for every new UTXO from. Since the enclave secret is used only once, such a construction is side-channel-free.

Least-privileged SGX. In $H_{SGX}$ and the examples above the fund can be spent by $pk_{SGX}$ alone, but it’s important to note that is not the only option. In fact, one can tune the knob between security and paralysis-tolerance to the best fit their needs. Specifically, for a desired level of paralysis-tolerance, one can design a DASP such that the SGX is least-privileged. For example, if the three shareholders only desire to tolerate up to one missing key share, what they can do is to move the funds into 3-out-of-4 multisig wallet where the 4th share is only known to the SGX enclave. If all of the parties are alive, then they can spend without use of the SGX node. If one of them is incapacitated, the enclave will release its share upon presentation of a Paralysis Proof. Therefore, even if the secret state of the SGX node (i.e., the fourth share) is leaked via a successful side-channel attack, the attacker cannot spend the fund unless two malicious parties collude. It can be shown that the SGX in the above DASP is least-privileged, in the sense that compromise of its secret state imparts minimal capabilities to an adversary. Intuitively, since we want to retain access even one player is incapacitated, the enclave must store a credential equivalent to that of the lost player. We leave formal specification of least-privileged SGXs for future work.

4.6 Paralysis Proofs via Ethereum

An Ethereum implementation of the ideal functionality $T_{DASS}$ is straightforward. Our reference implementation of a paralysis-free multisig wallet consists of 156 lines of commented Solidity code, and its main logic is shown in Figure 7.

This implementation differs from the ideal functionality only in minor engineering changes and optimizations. There is no way to asynchronously prune keyholders that fail to respond to a challenge in time in Ethereum, where all contract calls must be initiated by some user. We instead check and prune any signers that did not respond to a challenge at the beginning of each on-chain operation that requires checking or manipulating only valid signers. This ensures that the state of unparalyzed signers is correct, reflecting $L_t$ in $T_{DASS}$, before any contract action is processed.

A final caveat is that block timestamps are used to measure time; while this can be trivially replaced with block numbers, which are less susceptible to miner manipulation (timestamps are miner set), the bounded degree of manipulation and monotonically increasing

timestamp constraints on Ethereum provide some assurance that the timestamps are reasonably accurate for our purposes.

One useful property of the Ethereum-based realization is that the multisignature key holders need not necessarily run archival nodes: because a log is emitted whenever a user is accused, users can simply watch transaction receipts for an accusation against them, using any Ethereum full or lite client to respond by calling the respond function (guaranteed to work as long as an adversary cannot censor a user’s connection to the blockchain, given that the user accepts the relevant trust assumptions surrounding their choice of node software, hardware, and connectivity).

Full tested contract code, including logic for pruning incapacitated signers and updating the signature threshold, is published at [3].

5 Custodial Paralysis

Until now, this work has considered paralysis in a generalized setting—any setting in which multiple parties must authorize a transaction. Consequently, we have adopted a very narrow definition of paralysis as unavailability of a party expected to be online. Without defined behavior for users or valid transactions, there is no distinction between a party who honestly refuses to sign an illegitimate message and one who maliciously paralyzes a legitimate transaction. Unfortunately this makes it impossible to construct a mechanism that distinguishes between the two. In some settings, however, relationships between parties and their separation of responsibilities with regards to transaction signing are better defined.

In this section, we consider one such specific scenario and analyze the consequences of paralysis: we consider a centralized custodian, effectively a bank, responsible for providing access control to a user’s funds. Here the goal is not to downgrade access if the bank becomes unavailable or paralyzed, but rather to migrate to a completely distinct recovery policy when the custodian fails to perform its assigned duties or either. Unlike in the previous setting, keys are not equal in privilege and come with defined roles.

Why centralized custodians? Centralized service providers often offer better security than a user is capable of providing on their own, and can relieve users from the burdens of key management and storage. Centralized providers can also offer layered security services including sophisticated access control, two-factor authentication, account compromise detection, account recovery options, and account insurance. Adding to these benefits, such services are convenient in their availability and do not require their users to purchase custom or dedicated hardware. Ease of key management is a major consideration for many users when choosing so called “web wallets” [42] and such services have seen remarkable success with one US-based custodial service Coinbase at times claiming to control up to 10% of Bitcoins in active circulation [8].

Centralized custodians, however, are a single point of failure for both security and availability: they can steal funds directly or simply disappear. A number of such cases of fraudulent or otherwise insolvent services leading to losses for their users have been observed in practice [1, 74]. Such trust issues can be resolved by instead layering private key security, requiring the custodian to sign off
on all transactions in a 2-of-2 multisignature scheme. This protects the user from custodial theft, but does not ensure the user’s funds are available to spend. It does not ensure users funds cannot be paralyzed, either maliciously or accidentally, by the failure of the centralized custodian to sign transactions in a timely fashion.

Paralysis-proof Custodians In our setting, we assume a user entrusts a centralized custodian with either a key that directly controls their funds or is part of a multi-signature address. The custodian is responsible for authenticating the user before it authorizes any transaction with its key. Separately, the user stores a recovery key (e.g. on paper in a safety deposit box). Optimistically, the recovery key will never be required, as the custodian will remain available. If the recovery key had full control over the funds, then the custodian would offer little additional security and the funds would only be as safe as the recovery key. Instead, the recovery key is inert, controlling no funds, and can only obtain funds through a paralysis proof of the centralized custodian. If the custodian is paralyzed, control is migrated from the exchange to the recovery key.

While basic paralysis proofs would guard against unavailability of the custodian, they would not guard against a custodian that maliciously blocks legitimately authorized transactions. To resolve this, we extend the functionality of basic Paralysis Proofs to include a predicate that must be met for a player to issue a life signal. In our case, the predicate will ensure a life signal is issued only if the custodian faithfully attempted and failed to authenticate the user and transaction. The exact details of this predicate can vary depending on the authentication mechanism. We explore the set of custodian functionalities that are paralysis proof compatible.

Mechanisms for Authentication and Secondary Authentication To authenticate to the custodian, the client will need to participate in a possibly interactive authentication protocol with messages passed from the client to the custodian and potential back. As a simple example, a standard password authentication protocol would require the client to send the custodian the password. For a paralysis proof against a malicious custodian, the client needs to demonstrate that they did actually submit such a message and conversely, an honest custodian defending against a false paralysis proof needs to show no such message was sent or the message was invalid. Because the custodian is realized in an Trusted Execution Environment, we need not contend with the correctness of the messages themselves: we can rely on the TEE to produce correct messages and accept valid messages, so we simply need a mechanism for assured communication between the enclave and user.

Following the techniques of [40], any interactive authentication protocol can be realized by posting encrypted messages between the client and the custodian to the blockchain. A paralysis proof then is simply an on chain execution of the protocol where messages are delivered and logged via the blockchain. For example in the password authentication protocol, a paralysis proof can be realized by posting the password, encrypted under a key owned by the enclave, in a challenge transaction. In responding to the challenge, the enclave sees the included password and if and only if the password incorrect, will contest the paralysis proof. This can be applied to $n$ round protocols through repetition and extended to, for example, integrating federated authentication protocols or challenge response based second factor authentication mechanisms.

Third Party Authentication Another option is for the enclave to contact third parties directly to provide primary or a secondary authentication factor. Town Crier [91] demonstrated that it is possible to make a TLS request from an enclave to a third party service and condition behavior on the response. Combined with input from the user via the blockchain, this can be used to directly authenticate a user, to provide a second factor via services such as Authy or Twillo, or to ensure that a user still has an account with some service. Many of these mechanisms depend on trusting a third party, but in the case of established and widely used services this may be more palatable than trusting a custodian. Multiple services can of course be combined to further distribute trust at the cost of availability.

Account Recovery and Dual Paralysis Using the same mechanisms for interactive protocols or authenticating via third parties, the custodian can provide a paralysis proof compatible mechanism for account recovery in the case of lost credentials. Indeed, if the custodian relies on third parties for authentication, then it inherits the account recovery mechanism automatically. This is, of course, a double edged sword: the same mechanisms that are used for account recovery can be used to hijack the account. We are not limited, however, to simple password recovery mechanisms. If the custodian only controls one of two keys necessary to spend the funds, we can realize account recovery by requiring a paralysis proof against the user to migrate access control to the recovery key.

6 Smart Contract Proof-Of-Paralysis

Throughout this work we have explored the relationship between smart contracts and Paralysis Proofs, using, e.g., smart contracts to implement Paralysis Proofs in Section 4.6. In the broader smart contract community, it is well known that paralysis can occur within smart contracts themselves. A classic example are the two well-publicized Parity multisignature vulnerabilities [15] [76] which together froze hundreds of millions of US dollars in smart contracts. The Parity multisignature vulnerabilities are far from the only high-profile failures that resulted in paralysis; early analysis of Ethereum smart contract vulnerabilities [18] enumerated a number of vulnerable contracts with stuck funds and denial-of-service vulnerabilities. Some vulnerabilities are subtle, involving low-level platform details like the “gas” model for pricing computation [87].

Most smart contract vulnerabilities, despite their different manifestations, have a fundamental commonality: in each case, a smart contract was operating as intended until some unexpected change to the state of the contract, the network, or the computation model under which the contract was operating. These changes then caused subsequent executions of previously working code/contracts to fail, leaving the funds in a contract potentially “paralyzed,” and stuck indefinitely. This problem is so widespread and severe on Ethereum, that hard-fork-based manual remediation of affected contracts has been suggested as a major governance issue and debate [17] [33].

We find a natural solution to this class of stuck funds in software engineering tradition. When integrating various system components
An SGX-based solution can also leverage attestations and confidentiality. For example, static state $s_i$ might contain an initialization of a special testing account with some balance available to be transferred that is not present in the storage variables from the environment. Merkle proofs-of-inclusion for all state items containing availability of the contract (like the developers, for example Parity Technologies in [65]) are strongly incentivized to submit such proofs when they detect paralysis. Contracts desiring additional incentives can also incentivize a market of test executors by adding a substantial bounty upon successful execution of migrate.

The smart-contract based scheme has important advantages. Primarily, any user can prove paralysis using only on-chain data at any time, minimizing the trust surface required to just the code governing proof verification and recovery (and excluding complex attestation schemes and trust in enclave confidentiality). The scheme is also practical and optimistically efficient. We tested an example implementation provided at [48] of a Merkle-Patricia state item proof checker in a smart contract, usable for the $s_{\text{is_valid}}$ function reference in $\text{prog}_a$. In the optimistic case where funds are not paralyzed, our scheme adds no on-chain overhead or additional cost to contract operation. In the exceptional case of paralysis, our scheme’s cost is justified by the potential recovery of funds. Our initial tests suggested a cost of about 1.36 million gas per invocation of the functionality required by $s_{\text{is_valid}}$; this is approximately 1/6 of a full Ethereum block, or 12USD per storage location proof at the time of writing: expensive but not prohibitive.

Several issues are however present with this SGX-free smart-contract scheme. The on-chain Merkle proof verification is still somewhat costly, and a transaction/test can potentially access many storage locations. This may be acceptable in smart contract form, as tests can be potentially broken up into small/short pieces. This scheme financially incentivizes short tests, however, which may limit expressiveness of developers’ tests in an effort to make worst-case verification cost tractable in a volatile and unpredictable fee market. By making the cost of executing a test less dependent on the size of the test using SGX and off-chain computation, longer and more expressive tests become tractable. Such a scheme can optionally be operated in the Sealed-Glass Proof model in [80], enabling resilience against unbounded side channel leakage and removing confidentiality requirements, relying only on attestations for security.

An SGX-based solution can also leverage attestations and confidentiality. For example, if any off-chain or legacy systems are required in the integration test (e.g., when an oracle such as Town Crier [92] is used), they can be queried or emulated by SGX, or can use a trusted off-chain oracle. Also, confidential tests may be useful for some contracts. While unsuitable in a public network due to their ability to hide backdoors, one could imagine a contract between parties with neutrally-agreed-on third parties or arbiters responsible for maintaining independent anti-paralysis test suites.

### 7 Related Work

Bitcoin had built-in support for threshold signatures at launch, and access-structure scripts for Bitcoin have been discussed since at least 2012 (see, e.g., [67]). The witness for the built-in Bitcoin

![Figure 5: An example implementation of Paralysis Proofs for smart contracts.](image-url)
opcode is a list that consists of the individual signatures that meet
the threshold, which increases the on-chain verification complexity
(this is undesirable, cf. [49]).

Gennaro, Goldfeder and Narayanan[30] presented a novel ECDSA
threshold signature construction that reduces on-chain complexity
of multi-signature Bitcoin address control. However, this construc-
tion requires a rather complex setup using ZK proofs, and does not
support arbitrary access structures. Threshold Schnorr signatures
are far more efficient [77], with support planned for Bitcoin [88].

Mesh signatures [13] can be used to implement an arbitrary access
structure. Attribute-based signatures [50] is an alternative approach
that utilizes a trusted third party to implement arbitrary access
structures. These constructions rely on complex cryptographic
primitives such as bilinear pairings (which have no native support
in Bitcoin script language). By themselves, mesh signatures can
only support a static access structure (cf. Section 1).

Ethereum wallets such as Mist [56] and Gnosis [84] support multi-
signature access structures, along with other features such as daily
limits. However, these wallets are implemented via on-chain code,
which implies that users will incur higher costs when the complex-
ity of the access structure is greater. New versions of the Gnosis
wallet [55] allow for arbitrary, unrestricted challenge-based policy
migration, but do not formalize security or suggest secure policies.

Access-control policies with dynamic access-structures Secret
sharing schemes with revocation support do not provide the
same guarantees as a paralysis proof system, since such schemes
require actions by at least a threshold of players to update the access
structure (see [23, 90]). By contrast, a paralysis proof system enables
any player to remove incapacitated players. Privacy-preserving
cloud services can allow remote administrators to modify access-
control policies dynamically, via cryptographic constructions (see,
e.g., [35, 41]). Dynamic access-control policies for a non-confidential
cloud service may also benefit from dynamic access-control poli-
cies [59]. In all of these constructions, a policy modification affects
the ability of end-users to interact with the server, but the set of
administrators authorized to perform the modifications is static.

Credential-recovery Password systems allow recovery from sec-
ter loss, but require a trusted third party. See [12] for a survey.

Blockchains as censorship-resistant channels Bitcoin and other
cryptocurrencies have been proposed as a means to facilitate a
censorship-resistant channel. For instance, ZombieCoin [5] ana-
lyzes the prospects of a botnet command and control center using
Bitcoin. Another recent work [4] allows users behind a government
firewall to discover Tor entry nodes through a challenge-response
protocol with messages transmitted through cryptocurrency.

8 Paralysis Proof Systems Beyond Cryptocurrencies

The techniques we have introduced for Paralysis Proof Systems in
combining SGX with blockchains can be applied to settings other
than Paralysis Proofs and even to settings other than cryptocurrencies.
We give some examples here:

- **Daily spending limits:** It is possible to enforce limits on the
  amount of BTC that set of players can spend in a given interval
  of time. For example, players might be able to spend no more
  than 0.5 BTC per day. We explore this objective, and technical
  limitations in efficient solutions, in Appendix F.

- **Decryption:** The credentials controlled by a Paralysis Proof
  System need not be signing keys, but instead can be decryption
  keys. It is possible then, for example, to create a deadman’s
  switch. For example, a document can be decrypted by any of
  a set of journalists should its author be incapacitated.

- **Event-driven policies:** Using an oracle, e.g., [91], it is possible
to condition access-control policies on real-world events. For
example, daily spending limits might be denominated in USD
by accessing oracle feeds on exchange rates. Similarly, decryption
credentials for a document might be released for situations
other than incapacitation, e.g., if a document’s author is
prosecuted by a government. (This latter example would in all
likelihood require natural language processing, but this is not
beyond the capabilities of an enclave application.)

The last example involving prosecution does not require use of a
blockchain, of course. Many interesting SGX-enforceable access-
control policies do not. But use of a blockchain as a censorship-
resistant channel can help ensure that policies are enforced. For
example, release of a decryption key might be entangled with the
spending of cryptocurrency. A certain amount of cryptocurrency,
say, 10 BTC, might be spendable on condition that an oracle is
recently queried and the result consumed by an enclave application.
This approach provides an economic assurance of a censorship-
resistant channel from the blockchain to the enclave.

9 Conclusion

We have shown how Paralysis Proofs can enrich existing access-
control policies in a way that was previously unachievable without
a trusted third party. By leveraging Paralysis Proofs, DASSes allow
an access structure to be securely migrated—typically downgraded—
given the incapacitation of a player, the inability of a set of players
to act in concert, or the functional paralysis of a smart contract.

Our supporting formalism includes a formal DASP framework, and
security and functionality property definitions for DASSes and
DASPs, as well as UC-type ideal functionality for a DASS. Our ideal
functionality naturally suggests a proof sketch for security.

Paralysis Proofs and DASSes can be applied in many settings, and
we showcase three in the paper: cryptocurrency key loss, crypto-
currency custody failures, and smart contract failures, proposing
practical schemes for all three. We report on a simple DASS for
cryptocurrency key loss in Ethereum, and on a detailed exploration
concluding that DASS for Bitcoin is only practical using a TEE.

In summary, we believe that the combination of the advent of two
pivotal technologies, blockchains and trusted hardware (specifically
SGX), is a powerful one. It enables a powerful new range of access-
control regimes without the need for trusted third parties and, we
believe, will stimulate exploration of a broad spectrum of other
novel capabilities with applications beyond cryptocurrencies.
Acknowledgements

This work is funded in part by NSF grants CNS-1330599, CNS-1514163, CNS-1564102, and CNS-1704615, as well as ARO grant W911NF-16-1-0145.

References

[4] R3c3: Cryptographically-secure censorship resistant rendezvous using cryptographic techniques and strategies. In Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Centrality in the consensus computer. In Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Centrality in the consensus computer. In Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, in

Figure 6: Ideal blockchain. The entry point marked with + is only executed once. The parameter succ defines the validity of new items. A new item can only be appended to the evaluation if the success of succ outputs 1.

6. If succ(storage, inp) = 1 then
7. t = clock();
8. storage := storage | (t, P, inp);
9. output (receipt, inp);
10. else output (reject, inp)

---

A Addition Formalism

B Proof of Theorem 4.1

We recall Theorem 4.1:

Theorem 4.1 (The Security of \texttt{PSG}). Assume \texttt{SGX} ’s attestation scheme and the digital signature used in \texttt{PSG} are existentially unforgeable under chosen message attacks (EU-CMA). Then \texttt{PSG} UC-realizes \mathcal{F}_{\texttt{DASS}}[\mathcal{S}_p, \mathcal{R}, \mathcal{S}, \mathcal{M}] in the (\texttt{SGX}, \mathcal{F}_{\text{blockchain}})-hybrid model, for static adversaries.

Proof. For simplicity, we write \mathcal{F}_{\texttt{DASS}}[\mathcal{S}_p, \mathcal{R}, \mathcal{S}, \mathcal{M}] as \mathcal{F}_{\texttt{DASS}} from now on. To prove Theorem 4.1, it suffices to show that for the "dummy adversary" \mathcal{A}, there exists a PPT adversary Sim such that...
for any PPT environment $Z$.

$$\text{EXEC}_{\PiSGX, A, Z} \approx \text{EXEC}_{\mathcal{F}\text{DASS}, \text{Sim}, Z} \cdot$$  \hspace{1cm} (1)

Basically, the dummy adversary simply relays messages between the environment $Z$ and parties. In particular, $A$ corrupts parties when instructed by $Z$ and passes all gathered information to $Z$. We refer readers to Section 4.4.1 of [19] for details on emulation with respect to the dummy adversary.

We first present the construction of Sim, then we show that Sim satisfies Equation (1).

### B.1 Construction of Sim.

Sim generally proceeds as follows: if a message is sent by an honest party to $\mathcal{F}\text{DASS}$, Sim emulates the appropriate real-world “network traffic” for $Z$ using the information obtained from $\mathcal{F}\text{DASS}$. If a message is sent to $\mathcal{F}\text{blockchain}$ or $\PiSGX$ by a corrupted party, Sim intercepts the input and interact with $A$ with the help of $\mathcal{F}\text{DASS}$. We provide further details on the processing of specific messages.

**Initialization.** For an honest party $P_h$, Sim faithfully emulates $\PiSGX$ as if $P_h$ is called with a `init` message. In particular, Sim generates the key pair and simulates the initialization of $\PiSGX$ and $\mathcal{F}\text{blockchain}$ if not already initialized. If a malicious party $P_m$ sends `init` with corrupted parameters (i.e. different from those of $\mathcal{F}\text{DASS}$), Sim aborts after simulating an `abort` message to all parties.

**Access.** For an honest party $P_h$ calling $\mathcal{F}\text{DASS}$ with (access, inp), Sim computes a signature $\sigma$ over UTXOfund and inp (using the secret key generated in the initialization phase) and simulates $\PiSGX$ faithfully.

If $\PiSGX$ (simulated by Sim) is activated, Sim proceeds as follows: if a message is sent by an honest party $P_h$ to $\mathcal{F}\text{DASS}$, Sim emulates the appropriate real-world “network traffic” for $Z$ using the information obtained from $\mathcal{F}\text{DASS}$. If a message is sent to $\mathcal{F}\text{blockchain}$ or $\PiSGX$ by a corrupted party, Sim intercepts the input and interact with $A$ with the help of $\mathcal{F}\text{DASS}$. We provide further details on the processing of specific messages.

**Migration from $Sp$ to $Sp'$.** Without loss of generality, we only consider the cases where $P = P' \cup \{p_{k}\}$, namely the migrations that remove $P_k$ from the access structure. Other cases can be analyzed similarly.

If $\mathcal{F}\text{DASS}$ is activated by an honest party $P_h$ with input (`migrate, $P'$`) that removes $P_k$ from the current access structure, Sim proceeds as follows:

- If $P_k$ is honest or paralyzed, Sim emulates $P_i$’s part of $\PiSGX$ by computing $t_1$ and $t_2$. If $\mathcal{F}\text{DASS}$ authorizes the migration (indicated by a public $ok$ output), Sim delay the output for $\Delta$ time, put $t_2$ on the blockchain, and then permits $\mathcal{F}\text{DASS}$ to deliver the output to all parties. If $\mathcal{F}\text{DASS}$ rejects the migration, Sim spends the life signal on $P_k$’s behalf.

- If $P_i$ is malicious, Sim waits until $\mathcal{F}\text{DASS}$ asks whether $P_k$ chooses to pretend to be paralyzed. Sim generates $t_1$ and $t_2$ and sends $t_1$ to $\mathcal{F}\text{blockchain}$. If $A$ the output on behalf of $P_k$, Sim responds “yes” to $\mathcal{F}\text{DASS}$. If not, Sim sends “no” to $\mathcal{F}\text{DASS}$ and put $t_2$ to $\mathcal{F}\text{blockchain}$.

If $\PiSGX$ (simulated by Sim) is activated by a corrupted party $P_c$ with input (`migrate, UTXOfund, $P'$`) that removes $P_k$ from the current access structure, Sim computes $t_1$ and $t_2$ and send both to $A$ as if from $\mathcal{F}\text{blockchain}$ and then proceeds as follows:

- If $P_k$ is honest and alive: If $A$ sends $t_1$ to $\mathcal{F}\text{blockchain}$, Sim sends $t_1$ on $P_k$’s behalf, and send (`migrate, $sp'$`) to $\mathcal{F}\text{DASS}$ on $P_c$’s behalf at the same time.

- If $P_k$ is paralyzed: If $A$ sends both $t_1$ and $t_2$ to $\mathcal{F}\text{blockchain}$, Sim sends (`migrate, $sp'$`) to $\mathcal{F}\text{DASS}$ on $P_c$’s behalf.

- If $P_k$ is malicious: If $A$ sends both $t_1$ and $t_2$ to $\mathcal{F}\text{blockchain}$, and $A$ doesn’t send $t_1$, then Sim sends (`migrate, $sp'$`) to $\mathcal{F}\text{DASS}$ on $P_c$’s behalf, and sends “no” to $\mathcal{F}\text{DASS}$ when asked whether $P_k$ chooses to be paralyzed.

**B.2 Validity of Sim.**

We show that no environment can distinguish an interaction with $\mathcal{A}$ and $\PiSGX$ from one with Sim and $\mathcal{F}\text{DASS}$ by hybrid arguments. Consider a sequence of hybrids, starting with the real-world execution of $\PiSGX$. $H_1$ lets Sim to emulate $\PiSGX$ and $\mathcal{F}\text{blockchain}$. $H_2$ filters out the forgery attacks against $\SigmaSGX$ and $H_3$ filters out the forgery attacks against the signature scheme.

**Hybrid $H_1$.** Proceeds as in the real-world protocol, except that Sim emulates $\PiSGX$ and $\mathcal{F}\text{blockchain}$. Specially, Sim generates a key pair $(pk_{\text{att}}, sk_{\text{att}})$ for $\SigmaSGX$ and publishes $pk_{\text{att}}$. Whenever $A$ wants to communicate with $\PiSGX$, Sim records $A$’s messages and faithfully emulates $\PiSGX$’s behavior. Similarly, Sim emulates $\mathcal{F}\text{blockchain}$ by storing items internally.

As $A$’s view in $H_1$ is perfectly simulated as in the real world, no $Z$ can distinguish between $H_1$ and the real execution.

**Hybrid $H_2$.** Proceeds as in $H_1$, except for the following modifications. If $A$ invokes $\PiSGX$ with a correct message (`install, $\text{prog}_{\text{new}}$`), then for all sequential `resume` calls, Sim records a tuple (outp, $\sigma\text{SGX}$) where outp is the output of `prog_{new}` and $\sigma\text{SGX}$ is an attestation under $sk_{\text{att}}$. Let $\Omega$ denote the set of all such tuples. Whenever $A$ sends an attested output (outp, $\sigma\text{SGX}$) $\notin \Omega$ to Sim or an honest party, Sim aborts.

The indistinguishability between $H_1$ and $H_2$ can be shown by the following reduction to the the EU-CMA property of $\SigmaSGX$. In $H_1$, if $A$ sends forged attestations to Sim, signature verification will fail with all but negligible probability. If $Z$ can distinguish $H_2$ from $H_1$, $Z$ and $A$ can be used to win the game of signature forgery.

**Hybrid $H_3$.** Proceeds as in $H_2$, except for the following modifications. Suppose the set of public keys belonging to corrupted
parties is \( \{p_{ki}\}_{i=1}^{N} \). If \( A \) sends \((\text{spend}, \sigma, \ldots)\) and \( \sigma \) verifies under a public key \( p_k \notin \{p_{ki}\}_{i=1}^{N} \), Sim aborts.

Similarly, the indistinguishability between \( H_2 \) and \( H_3 \) can be shown by a reduction to the EU-CMA property of signature scheme.

It remains to observe that \( H_3 \) is identical to the ideal protocol with Sim. \( \square \)

C Additional Code Sample

C.1 Paralysis Proofs via Ethereum

Figure 7 shows the code snippet of the Ethereum implementation of Paralysis Proofs.

D Paralysis Proofs via Covenants

In the context of Bitcoin scripts, the notion of a covenant allows to put restrictions on the way that an output can be spent. Covenants were introduced by Moser, Eyal and Sirer [58], following an early idea by Maxwell [53]. Another generic method for covenants was given by O’Connor and Piekarska [61], and an efficient implementation of covenants (\( \text{OP\_PUSHTXDATA} \)) was created by Lau [45]. So far, covenants support has not been enabled on the Bitcoin mainnet.

The mechanism of [58] supports a recursive covenant by letting the interpreter replace the Pattern keyword with the covenant itself. The recursion is required in our paralysis use-case, because the funds must be restored back to the original covenant whenever an accusation attempt fails. However, the single Pattern capability of [58] is inadequate for the paralysis covenant, because we wish to move the funds between different covenants that depend on the subset of remaining shareholders. Fortunately, the implementation of [45] supports multiple recursive patterns, by hashing fixed and variable data and then comparing the result to the output P2SH address [7] or the SegWit P2WPKH/P2WSH [44] (as well as Merkelized syntax trees [43, 70]).

An exemplary covenant pattern is illustrated in Figure 8, using syntax that is similar to that of [58]. In this example, three shareholders can spend the entire amount \( V \) with no restrictions, by using the 3-out-of-3 multisig condition of the Pattern123 covenant. Any two shareholders can accuse the third shareholder of being paralyzed, by moving the entire fund of \( V \) coins into an Pattern1J covenant that lets them spend the coins after a relative timeout of 150 blocks. While the 150 blocks are still being created, the third shareholder can move the funds back into the initial covenant Pattern123. Similarly, any single shareholder can accuse the two other shareholders of being paralyzed, by moving the \( V \) coins into the Pattern1 covenant.

Note that the covenants Pattern1J and Pattern1 must be distinct for different values of \( I, J \), in order to avoid collusion attacks. For example, if Pattern1J allowed any 2-out-of-3 to spend the \( V \) funds after the timeout, then two malicious shareholders \( P_2, P_3 \) could pretend that \( P_3 \) is paralysed, so that \( P_1, P_2 \) would accuse \( P_3 \), and after the 150 blocks timeout \( P_3 \) will spend the funds arbitrarily (without the consent of the honest \( P_i \)).

There is certain similarity between the SGX protocol of Figure 2 and the covenants implementation of Figure 8. The main difference is that the \( p_{k1}, p_{k2} \) multisig replaces \( p_{SGX} \) in the condition \( (p_{SGX} \land CSV \geq \Delta) \). Hence, by taking the paralysis use-case as an example, it can be inferred that the complexity of the covenants approach is significantly higher than that of an SGX implementation (in terms of conceptual as well as on-chain complexity, see also Appendix E). As there have been recent proposals to support stateless covenants in Ethereum (for better scalability, cf. [83]), the comparative advantages of our SGX-based design may prove useful in other contexts too.
The ideal functionality of Section 4 is threshold predicates that verification of such schemes will be impractical or costly.

The Complexity of Access Structure Realizations

The ideal functionality of Section 4 can therefore be replaced with an extended functionality that supports an access structure, and the Bitcoin protocol of Figure 2 will essentially remain the same. This is because the off-chain complexity of creating the signature will be handled by the SGX enclave code, and the on-chain complexity will be absorbed into a verification against pk_{SGX}.

It is worth considering whether it is inherently that case that the high efficiency requires SGX, or whether it is possible to design a cryptocurrency with built-in support to access structure based signatures. In fact, certain support is offered via the use of Merkliized Abstract Syntax Trees [43, 70] and Schnorr aggregate signatures [88]. As in the "Large multi-signature constructs" of [43], we can for example have a Merkle tree with $2^{18}$ leaves, such that all but two of the leaf nodes require a multisig by a specific subset of $\{P_1, P_2, \ldots, P_{40}\}$ of size 35 (excluding subsets that already include the privileged sets $\{P_1, P_2, P_3, P_4\}$ and $\{P_1, P_35, P_36, P_37, P_38, P_39, P_40\}$, without double counting), put only the root hash on the blockchain, and expect a valid Merkle authentication path to spend the coins. Further, the script of the leaf can use a single aggregated public key that is created from the public keys of the 35 signers (using delinearization [9, 82]), so that the on-chain complexity is on part with that of verifying one ordinary signature. Regarding the total on-chain complexity, we have that transaction that spends the funds consists of one aggregated signature for the leaf node and a Merkle authentication path of 18 sibling hashes.

However, per the discussion of OP_EVAL in [43], the use of a Merklized Abstract Syntax Trees becomes significantly more challenging for a predicate that involves a more complex relation than a logical OR among the leaves. For instance, if the access structure specified that $P_1, P_2, P_3$ must consent, and either $P_4, P_5$ or $P_6, P_7$ must also consent, then this cannot be handled by the implementation of [43]. By contrast, SGX can handle this instance just as easily as the previous example.

As the above discussion illustrates, harnessing SGX to spend funds according to an access structure can be highly useful even for a cryptocurrency with a Turing-complete scripting language (such as Ethereum). Let us point out that as long as [88] is not yet operational, it can be quite beneficial to employ SGX even for threshold signatures, since an ECDSA threshold scheme (without a trusted dealer) is rather complex, cf. [30].

The use of access structures in a cryptocurrency can also incorporate a notion of time, which in turn can help to avoid system paralysis that is caused by disagreement together with the disappearance of some players. For instance, the functionality can require 75% of the active players to agree on how to spend the funds, but 20% after three years. In the UTXO model of Bitcoin, this can be accomplished via trusted hardware: whenever the players agree to spend the funds they will specify absolute timeouts for the 50% and 20% cutoffs (using CLTV [79]), and whenever the SGX enclave is asked to remove an incapacitated player it will create a transaction whose
output hardcodes the same absolute timeouts as the input that is being spent. If the access structures for the different points in time are complex, the trusted hardware based implementation will be particularly beneficial (otherwise covenants could be used).

F Daily Withdrawal Limit using SGX

Let us consider a functionality $\mathcal{F}_{\text{daily}}$ that allows $N$ shareholders to spend the funds if at least $\mu N$ of them reach an agreement (for $\mu \geq 1$), and allows each individual shareholder to spend a small portion of the funds (e.g., 0.1%) each day. Moreover, the functionality allows $\rho N$ shareholders to disable the daily spending of funds by individual shareholders (for $\rho \leq \mu$, which is useful in the case that some shareholders appear to spend too much). By using $\rho < \mu$, it is easier to block the daily withdrawals than to reach consensus on a large expenditure.

It may be quite useful to combine $\mathcal{F}_{\text{daily}}$ with a functionality for paralysis proofs, but for simplicity we focus in this section on a bare implementation of just $\mathcal{F}_{\text{daily}}$ itself. Given an expressive enough covenants support for Bitcoin (such as [45]), it is possible — though quite complex — to implement $\mathcal{F}_{\text{daily}}$ using similar methods to the ones that we describe in Appendix D. However, let us present here the more efficient implementation that relies on trusted hardware, and can be deployed on the current Bitcoin mainnet.

The SGX-based protocol $\Pi_{\text{daily}}$ that implements $\mathcal{F}_{\text{daily}}$ is given in Figure 9.

Since Bitcoin outputs must be fully consumed, $\Pi_{\text{daily}}$ does not realize $\mathcal{F}_{\text{daily}}$ exactly, but instead lets each one of the shareholders perform a daily withdrawal, in sequential order. Thus, the first shareholder has the privilege to withdraw a small amount on the first day, the second shareholder can withdraw a small amount on the second day, the third shareholder on the third day, and so on. If for example the third shareholder did not withdraw, then on the fourth day any single shareholder can withdraw a small amount (on a first come first served basis), but on the fifth day the sequential order resumes and the forth shareholder will have the privilege to withdraw.

It should be noted that in a cryptocurrency that uses the accounts model rather than the UTXO model (e.g., Ethereum), a more expressive realization of $\mathcal{F}_{\text{daily}}$ is possible. E.g., multiple shareholders can withdraw small amounts as long as the daily limit has not yet been reached.

The gist of $\Pi_{\text{daily}}$ is an embedding of a public key $pk_{\text{SGX}}$, into the spending transaction, corresponding to the shareholder $P_i$ who currently has the daily withdrawal privilege. Since the secret key $sk_{\text{SGX}}$, is known only to the SGX enclave, $P_i$ cannot spend the funds arbitrarily, but instead has to submit a request to spend a small amount $V'$ of the $V$ coins to an arbitrary destination $T'$. The enclave will thus also produce a new output for the rest of the $V - V'$ funds, with $pk_{\text{SGX}}$, embedded into it.

Since $P_i$ may not necessarily wish to withdraw, the output that the enclaves produces also allows spending of a small amount with a special master public key $pk_{\text{SGX}}$, but only after a relative timeout of $\Delta$ blocks (since Bitcoin blocks are created once every 10 minutes on average, $\Delta = 144$ blocks implies $\approx 1$ day). Hence, any shareholder who submitted a request to withdraw from the current funds will be able to spend the signed transaction that the enclave produced for her, but only after $\Delta$ blocks so that $P_i$ has the opportunity to spend first.

In case $\mu N$ shareholders wish to spend an arbitrary amount, or in case $\rho N$ shareholders wish to disable the daily withdrawal feature, they can submit their $\mu N$ (or $\rho N$) signatures to the enclave and receive a signed transaction that takes precedence over any daily withdrawal transaction. This is accomplished by using a small relative timeout $\Delta'$ in the condition that allows the current privileged shareholder to perform a daily withdrawal, so that the transaction that was agreed upon by $\mu N$ (or $\rho N$) shareholders can be incorporated into the blockchain earlier (e.g., $\Delta' = 3$ is reasonable).

Other parts of the $\Pi_{\text{daily}}$ protocol (in particular the setup procedure) are identical to $\Pi_{\text{SGX}}$, see Section 4 for details.

G Purely Bitcoin-Based Paralysis Proofs

A Paralysis Proof mechanism can also be implemented without SGX (on the current Bitcoin mainnet), albeit with subpar security and more than exponential overhead.
Each player $P_i$ will prepare unsigned transactions $\{t_{i,j,k}\}_{j\in[N], k\in[K]}$. All players will sign transactions $t'_{i,j,k}$ that take UTXO$_0$ and the output of $t_{i,j,k}$ as inputs (these transactions are similar to $t_1$). $K$ is a security parameter specifying the number of accusation attempts that can be made. Figure 10 illustrates the transactions in the aforementioned scheme.

This scheme can be implemented post-SegWit [26], where transaction hash (txid) excludes the ScriptSig witness. In particular, SegWit allows one to prepare $t'_{i,j,k}$ and condition its validity on that of unsigned $t_{i,j,k}$.

After every player receives all the signed transactions, the players will move the high-value fund into UTXO$_0$. This guarantees atomicity: either every player will have the ability to eliminate all the incapacitated player, or none of the player will have this ability. The output of $t_{i,j,k}$ requires a signature from $P_j$ before the CSV timeout and a signature from $P_i$ after the CSV timeout, and $P_j$ may embed this signature into $t'_{i,j,k}$ after $P_j$ failed to spend the output of $t_{i,j,k}$ on the blockchain. Since UTXO$_0$ requires the signatures of all parties, the only way to eliminate an incapacitated player is by using the signed transactions $t'_{i,j,k}$ that were prepared in advance.

The parameter $K$ specifies the number of accusation attempts that can be made; hence a malicious player that pretends to be incapacitated more than $K$ times will break this scheme. The SGX scheme does not exhibit this deficiency, because any player can send a fresh small amount of bitcoins to the enclave and thereby create an accusation transaction.

Furthermore, in order to support sequences of $\ell > 1$ incapacitated players, the $N$ players will need to prepare in advance additional transactions that spend the outputs of $t'_{i,j,k}$ in order to eliminate another player, and so on. The scheme offers the most safety when $\ell = N - 1$, as this implies that any lone active player (i.e., all other players became incapacitated) will be able to gain control over the fund. The number of signed transactions that need to be prepared in advance is

$$f(\ell, N, K) \triangleq KN(N-1)K(N-1)(N-2)\cdots K(N-\ell+1)(N-\ell) \geq \Omega(K^\ell N^\ell).$$

Thus, $\ell = N - 1$ implies that $f(\ell, N, K)$ grows faster than $g(N) = 2^N$. 

Figure 10: Bitcoin-based Paralysis Proofs with $N$ players (with public keys $\{pk_n\}_{n\in[N]}$). Each player $P_i$ will prepare unsigned transactions $\{t_{i,j,k}\}_{j\in[N], k\in[K]}$. All players will sign transactions $t'_{i,j,k}$. 

\[
\begin{align*}
t_{i,j,k}: & \quad 0.00001 \text{ BTC} \quad pk_i \\
& \quad 0.00001 \text{ BTC} \quad pk_j \lor (pk_i \land (\text{CSV} \geq 144)) \\
\end{align*}
\]

\[
\begin{align*}
t'_{i,j,k}: & \quad 5000 \text{ BTC} \quad \land_{n\in[N]} pk_n \\
& \quad 5000 \text{ BTC} \quad \land_{n\neq j} pk_n \\
\end{align*}
\]