

# MantaPay Protocol Specification

## v1.0.0

Shumo Chu, Boyuan Feng, Brandon H. Gomes, and Todd Norton \*

May 11, 2022

### Abstract

MantaPay is an implementation of a *decentralized anonymous payment* scheme based on the MANTADAP protocol outlined in the original [MANTA whitepaper](#).

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Notation</b>	<b>2</b>
<b>3</b>	<b>Concepts</b>	<b>2</b>
3.1	Assets	2
3.2	Addresses	3
3.3	Ledger	3
3.3.1	UTXOs and the UTXOSet	4
3.3.2	EncryptedNotes	4
3.3.3	VoidNumbers and the VoidNumberSet	4
<b>4</b>	<b>Abstract Protocol</b>	<b>5</b>
4.1	Abstract Cryptographic Schemes	5
4.1.1	Hash Function	5
4.1.2	Commitment Scheme	5
4.1.3	Key-Derivation Function	5
4.1.4	Randomizable Key-Derivation Function	5
4.1.5	Key-Agreement Scheme	6
4.1.6	Message Authentication Code	6
4.1.7	Signature Scheme	6
4.1.8	Symmetric-Key Encryption Scheme	6
4.1.9	Message Digest Cipher	7
4.1.10	Hybrid Public Key Encryption Scheme	7
4.1.11	Authenticated Hybrid Public Key Encryption Scheme	8
4.1.12	Dynamic Cryptographic Accumulator	8
4.1.13	Non-Interactive Zero-Knowledge Proving System	8
4.2	Addresses and Key Components	9
4.3	Transfer Protocol	11
4.4	Batched Transactions	15
<b>5</b>	<b>Concrete Protocol</b>	<b>16</b>
5.1	Poseidon Permutation	16
5.2	Elliptic Curve Cryptography	16
5.3	Concrete Cryptographic Schemes	16
5.4	AssetValue Bounds Check	17
<b>6</b>	<b>Acknowledgements</b>	<b>17</b>
<b>7</b>	<b>References</b>	<b>17</b>

---

\*ordered alphabetically

# 1 Introduction

MantaPay aims to solve the long-standing privacy problems facing cryptocurrencies in the Web3 age. At its heart, it uses various cryptographic constructions including NIZK (non-interactive zero knowledge proof) systems to ensure user privacy from *first principles*.

Protocol	Cryptographic Primitives	Consensus	Layer	Multi-Asset
ZCash (Sapling)	NIZK	PoW	1	<b>✗</b>
Monero	RingCT/NIZK	PoW	1	<b>✗</b>
Tornado Cash (Nova)	NIZK	<b>✗</b>	2	<b>✓</b>
MantaPay 1.0.0	NIZK	PoS	1	<b>✓</b>

**Table 1:** Comparison of MantaPay with previous constructions

## 2 Notation

The following notation is used throughout this specification:

- **Type** is the type of types<sup>1</sup>.
- If  $x : T$  then  $x$  is a value and  $T$  is a type, denoted  $T : \text{Type}$ , and we say that  $x$  *has type*  $T$ .
- **Bool** is the type of booleans with values **True** and **False**.
- For any types  $A : \text{Type}$  and  $B : \text{Type}$  we denote the *type of functions* from  $A$  to  $B$  as  $A \rightarrow B : \text{Type}$ .
- For any types  $A : \text{Type}$  and  $B : \text{Type}$  we denote the *product type* over  $A$  and  $B$  as  $A \times B : \text{Type}$  with constructor  $(-, -) : A \rightarrow (B \rightarrow A \times B)$ . Depending on context, we may omit the constructor and inline the pair into another constructor/destructor. For example, if  $f : A \times B \rightarrow C$  we can denote  $f((a, b))$  as  $f(a, b)$  to reduce the number of parentheses.
- For any type  $T : \text{Type}$ , we define  $\text{Option}(T) : \text{Type}$  as the inductive type with constructors:

$$\begin{aligned} \text{None} &: \text{Option}(T) \\ \text{Some} &: T \rightarrow \text{Option}(T) \end{aligned}$$

- We denote the *type of finite sets* over a type  $T : \text{Type}$  as  $\text{FinSet}(T) : \text{Type}$ . The membership predicate for a value  $x : T$  in a finite set  $S : \text{FinSet}(T)$  is denoted  $x \in S$ .
- We denote the *type of finite ordered sets* over a type  $T : \text{Type}$  as  $\text{List}(T) : \text{Type}$ . This can either be defined by an inductive type or as a  $\text{FinSet}(T)$  with a fixed ordering. We denote the constructor for a list as  $[\dots]$  for an arbitrary set of elements.
- We denote the *type of distributions* over a type  $T : \text{Type}$  as  $\mathfrak{D}(T) : \text{Type}$ . A value  $x$  sampled from  $\mathfrak{D}(T)$  is denoted  $x \sim \mathfrak{D}(T)$  and the fact that the value  $x$  belongs to the range of  $\mathfrak{D}(T)$  is denoted  $x \in \mathfrak{D}(T)$ . So namely,  $y \in \{x \mid x \sim \mathfrak{D}(T)\} \leftrightarrow y \in \mathfrak{D}(T)$ .
- We denote the equality predicate as  $(- = -) : T \times T \rightarrow \text{Type}$  and the equality function as  $\text{eq} : T \times T \rightarrow \text{Bool}$  whenever they exist.
- Depending on the context, the notation  $|\cdot|$  denotes either the absolute value of a quantity, the length of a list, the number of characters in a string, or the cardinality of a set.

## 3 Concepts

### 3.1 Assets

The **Asset** is the fundamental currency object in the MantaPay protocol. An asset  $a : \text{Asset}$  is a tuple

$$a = (a.\text{id}, a.\text{value}) : \text{AssetId} \times \text{AssetValue}$$

<sup>1</sup>By *type of types*, we mean the type of *first-level* types in some family of type universes. Discussion of the type theory necessary to make these notions rigorous is beyond the scope of this paper.

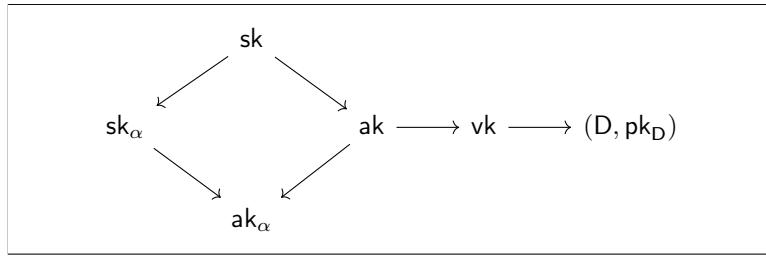
where the `AssetId` encodes the type of currency stored in  $a$  and the `AssetValue` encodes how many units of that currency are stored in  $a$ . **MantaPay** is a *decentralized anonymous payment* protocol which facilitates the private ownership and private transfer of **Asset** objects.

Whenever an **Asset** is being used in a public setting, we simply refer to it as an **Asset**, but when the `AssetId` and/or `AssetValue` of a particular **Asset** is meant to be hidden from public view, we refer to the **Asset** as either, *secret*, *private*, *hidden*, or *shielded*.

**Assets** are the basic building-blocks of *transactions* which consume a set of input **Assets** and produce a set of transformed output **Assets**. To preserve the economic value stored in **Assets**, the sum of the input `AssetValues` must balance the sum of the output `AssetValues`, and all assets in a single transaction must have the same `AssetId`<sup>2</sup>. This is called a *balanced transfer*: no `AssetValue` is created or destroyed in the process. The MantaPay protocol uses a distributed algorithm called **Transfer** to perform balanced transfers and ensure that they are valid.

### 3.2 Addresses

In order for MantaPay participants to receive **Assets** via the **Transfer** protocol, they create a *shielded addresses* which they use as identifiers to represent them on the ledger.



**Figure 1:** Key Schedule for MantaPay.

MantaPay uses four kinds of keys all derived from a base secret key  $sk$ , which give the following kinds of privileged access in the protocol:

- **Shielded Address** (send): Access to the shielded address  $(D, pk_D)$  gives the user the right to send **Assets** to the owner of the associated  $sk$ . The diversifier  $D$  allows the owner of a given  $sk$  key to generate many shielded addresses with the same backing spend authority.
- **Viewing Key** (view): Access to the viewing key  $vk$  gives the user the right to view all transactions for the owner of the associated  $sk$ .
- **Proof Authorization Key** (prove): Access to proof authorization key  $ak$  gives the user the right to build the transaction proof on behalf of the owner of  $sk$ . In order to delegate proof generation to another party, the owner of the secret key generates a randomizer  $\alpha$  and sends it to the prover which generates the proof. The owner then signs the transaction against  $ak_\alpha$  with their randomized key  $sk_\alpha$  which proves that they have knowledge of  $sk$ .
- **Spending Key** (spend): Access to the spending key  $sk$  gives total control over the assets owned by this secret, including spending, proof generation, and viewing.

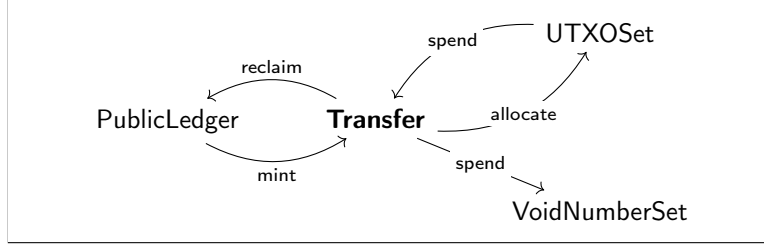
Participants in MantaPay are represented by their addresses, but they are not unique representations, since one participant may have access to more than one secret key. See § 4.2 for more information on how these keys are constructed and used for spending, proving, viewing, and receiving.

### 3.3 Ledger

Preserving the economic value of **Assets** requires more than just balanced transfers. It also requires that **Assets** are owned by exactly one address at a time, namely, that the ability to spend an **Asset** can be proved before a transfer and revoked after a transfer. It is not simply the *information-content* of an **Asset** that should be transferred, but the *ability to spend the asset in the future*, which should be transferred. To enforce this second invariant we can use a public ledger<sup>3</sup> that keeps track of the movement of **Assets** from one participant

<sup>2</sup>It is beyond the scope of this paper to discuss transactions with inputs and outputs that feature different `AssetIds`, like those that would be featured in a *decentralized anonymous exchange*.

to another. Unfortunately, using a public ledger alone does not allow participants to remain anonymous, so MantaPay extends the public ledger by adding a special account called the *shielded asset pool* which is responsible for keeping track of the Assets which have been anonymized by the protocol. We denote the three ledger types in the protocol as follows: the public ledger as **PublicLedger**, the shielded asset pool as **ShieldedAssetPool**, and the combined ledger we denote **Ledger**.



**Figure 2:** Life cycle of an Asset.

The **ShieldedAssetPool** is made up of three parts that are used to enforce the balanced transfer of Assets among anonymous participants:

1. § 3.3.1 **UTXOSet**: The **UTXOSet** is a collection of ownership claims to subsets of the **ShieldedAssetPool** (called **UTXOs**), each one referring to an allocated Asset transferred to a participant of the protocol.
2. § 3.3.2 **EncryptedNotes**: For every **UTXO** there is a matching **EncryptedNote** which contains information necessary to spend the Asset, which can be used to *provably reconstruct* the **UTXO** convincing the **Ledger** of unique ownership. The **EncryptedNote** can only be decrypted by the recipient of the Asset or the designated viewer of the **UTXO**, specifically, the correct viewing key *vk*. See § 3.2 for more.
3. § 3.3.3 **VoidNumberSet**: The **VoidNumberSet** is a collection of commitments, like **UTXOs**, but which track the *spent state* of an Asset and are used to prove to the **Ledger** that an Asset is spent *exactly one time*.

The operation of these different parts of the **ShieldedAssetPool** is elaborated in the following subsections.

### 3.3.1 UTXOs and the UTXOSet

An *unspent transaction output*, or **UTXO** for short, represents a claim to the output of a balanced transfer which has otherwise *not yet been spent*. Every balanced transfer can produce some number of *public outputs*, represented by Assets, and/or *private outputs*, represented by **UTXOs**, and these **UTXOs** are stored in the **UTXOSet** of the **ShieldedAssetPool**. A **UTXO** can only be claimed by the participant who owns the underlying Asset, where ownership means *knowledge of the correct spending key* and the **Transfer** protocol requires that all inputs to a balanced transfer *prove* that they own a **UTXO** which the **ShieldedAssetPool** has already seen in the past. The **UTXOSet** is *append-only* since it represents the past state of *unspent* Assets. **UTXOs** can only be added to the **UTXOSet** as outputs in the execution of a **Transfer** which the **Ledger** checks for correctness.

### 3.3.2 EncryptedNotes

In order to find out what Asset a **UTXO** is connected to, every **UTXO** comes with an associated **EncryptedNote** which stores two pieces of information, the underlying Asset, and an ephemeral public key, a value which allows the new owner of the Asset to reconstruct the **UTXO**. Being able to *provably reconstruct* a correct **UTXO** is a prerequisite to ownership and the ability to spend the Asset in the future. Once a participant spends an Asset that they can decrypt, they build a new **EncryptedNote** for the next participant that they sent their Assets to, so that they can then spend it, and so on. This is called the *in-band secret distribution*.

### 3.3.3 VoidNumbers and the VoidNumberSet

Once the ability to spend an Asset is extracted from a (**UTXO**, **EncryptedNote**) pair, the **ShieldedAssetPool** requires another commitment in order to spend the Asset, transferring it to another participant. This commitment, called the **VoidNumber**, represents the revocation of the right to spend the Asset in the future, and ensures that the same Asset cannot be spent twice. Like the **UTXOSet**, the **VoidNumberSet** is *append-only* since it represents the past state of *spent* Assets. **VoidNumbers** can only be added to the **VoidNumberSet** as inputs in the execution of a **Transfer** which the **Ledger** checks for correctness.

<sup>3</sup>A public (or private) ledger is not enough to solve the *provable-ownership problem* or the *double-spending problem*. A *consensus mechanism* is also required to ensure that all participants agree on the current state and state transformations of the ledger. The design and specification of the consensus mechanism that secures the MantaPay ledger is beyond the scope of this paper.

## 4 Abstract Protocol

### 4.1 Abstract Cryptographic Schemes

In the following section, we outline the formal specifications for all of the *cryptographic schemes* used in the MantaPay protocol.

**Definition 4.1.1** (Hash Function). A *hash function* HASH is defined by the schema:

$$\begin{aligned} \text{Input} &: \text{Type} \\ \text{Output} &: \text{Type} \\ \text{hash} &: \text{Input} \rightarrow \text{Output} \end{aligned}$$

with the following properties:

- **Collision Resistance:** It is infeasible to find  $a, b : \text{Input}$  such that  $a \neq b$  and  $\text{hash}(a) = \text{hash}(b)$ .
- **Pre-Image Resistance:** Given  $y : \text{Output}$ , it is infeasible to find an  $x : \text{Input}$  such that  $\text{hash}(x) = y$ .
- **Second Pre-Image Resistance:** Given  $a : \text{Input}$ , it is infeasible to find another  $b : \text{Input}$  such that  $a \neq b$  and  $\text{hash}(a) = \text{hash}(b)$ .

We can also ask that a hash function be *binding* or *hiding* as in the above *Commitment Scheme* definition if we partition the *Input* space into a separate *Randomness* and *Input* space.

**Notation:** For convenience, we may refer to  $\text{HASH.hash}(x)$  by  $\text{HASH}(x)$ .

**Definition 4.1.2** (Commitment Scheme). A *commitment scheme* COM is defined by the schema:

$$\begin{aligned} \text{Input} &: \text{Type} \\ \text{Output} &: \text{Type} \\ \text{Randomness} &: \text{Type} \\ \text{RandomnessDistribution} &: \mathcal{D}(\text{Randomness}) \\ \text{commit} &: \text{Randomness} \times \text{Input} \rightarrow \text{Output} \end{aligned}$$

with the following properties:

- **Binding:** It is infeasible to find an  $x, y : \text{Input}$  and  $r, s : \text{Randomness}$  such that  $x \neq y$  and  $\text{commit}(r, x) = \text{commit}(s, y)$ .
- **Hiding:** For all  $x, y : \text{Input}$ , the distributions  $\{\text{commit}(r, x) \mid r \sim \text{RandomnessDistribution}\}$  and  $\{\text{commit}(r, y) \mid r \sim \text{RandomnessDistribution}\}$  are *computationally indistinguishable*.

**Notation:** For convenience, we may refer to  $\text{COM.commit}(r, x)$  by  $\text{COM}_r(x)$ .

**Definition 4.1.3** (Key-Derivation Function). A *key-derivation function* KDF is defined by the schema:

$$\begin{aligned} \text{Input} &: \text{Type} \\ \text{Output} &: \text{Type} \\ \text{derive} &: \text{Input} \rightarrow \text{Output} \end{aligned}$$

**Notation:** For convenience, we may refer to  $\text{KDF.derive}(x)$  by  $\text{KDF}(x)$ .

**NB:** This abstract definition covers many different cases of key related functions. The security properties of a specific KDF are outlined wherever it's used.

**Definition 4.1.4** (Randomizable Key-Derivation Function). A *randomizable key derivation function* rKDF extends a KDF scheme by adding the following randomization:

$$\begin{aligned} \text{Randomness} &: \text{Type} \\ \text{RandomnessDistribution} &: \mathcal{D}(\text{Randomness}) \\ \text{rand}^I &: \text{Randomness} \times \text{Input} \rightarrow \text{Input} \\ \text{rand}^O &: \text{Randomness} \times \text{Output} \rightarrow \text{Output} \end{aligned}$$

where  $\text{rand}$  can be denoted without the type marker if the context is clear. The scheme has the following properties:

- **Random Derivation:** For all  $x : \text{Input}$  and  $\alpha : \text{Randomness}$ , we have  $\text{rand}(\alpha, \text{derive}(x)) = \text{derive}(\text{rand}(\alpha, x))$ .

**Notation:** For convenience, we may refer to  $\text{rand}(\alpha, x)$  by  $\text{rand}_\alpha(x)$  and refer to  $\text{rand}_\alpha(\text{derive}(x))$  by  $\text{derive}_\alpha(x)$ .

**Definition 4.1.5** (Key-Agreement Scheme). A *key-agreement scheme* KA is defined by the schema:

$\text{SecretKey} : \text{Type}$   
 $\text{PublicKey} : \text{Type}$   
 $\text{SharedSecret} : \text{Type}$   
 $\text{SecretKeyDistribution} : \mathcal{D}(\text{SecretKey})$   
 $\text{derive} : \text{SecretKey} \rightarrow \text{PublicKey}$   
 $\text{agree} : \text{SecretKey} \times \text{PublicKey} \rightarrow \text{SharedSecret}$

with the following properties:

- **Agreement:** For all  $\text{sk}_1, \text{sk}_2 : \text{SecretKey}$ ,  $\text{agree}(\text{sk}_1, \text{derive}(\text{sk}_2)) = \text{agree}(\text{sk}_2, \text{derive}(\text{sk}_1))$
- **Passive Security:** Even if an adversary eavesdrops on the network communication, she cannot forge the agreed secret unless she knows how to find a preimage for  $\text{derive}$  which should be as hard as a known hard cryptography problem like the Diffie-Hellman Problem.
- **Known-key Security:** Suppose an adversary learned a shared secret from a past session, then, the adversary does not gain any additional information by combining the past key and public visible data for the purpose of deducing future shared secrets.
- **No Key Control:** The shared secrets are determined by both parties, neither party can control the outcome of the shared secret by restricting it to lie in some predetermined small set.

**Definition 4.1.6** (Message Authentication Code). A *message authentication code* MAC is given by the schema:

$\text{SecretKey} : \text{Type}$   
 $\text{Message} : \text{Type}$   
 $\text{Tag} : \text{Type}$   
 $\text{sign} : \text{SecretKey} \times \text{Message} \rightarrow \text{Tag}$   
 $\text{verify} : \text{SecretKey} \times \text{Message} \times \text{Tag} \rightarrow \text{Bool}$

with the following properties:

- **Completeness:** For all  $\text{sk} : \text{SecretKey}$ ,  $m : \text{Message}$ , we have that  $\text{verify}(\text{sk}, m, \text{sign}(\text{sk}, m)) = \text{True}$ .
- **Unforgeability:** For any key  $\text{sk} : \text{SecretKey}$  and efficient adversary  $\mathcal{A}_{\text{sk}}$  with oracle access to  $\text{sign}(\text{sk}, -)$  the following probability is negligible:

$$\Pr \left[ \text{verify}(\text{sk}, m, t) = \text{True} \mid \begin{array}{l} (m, t) \sim \mathcal{A}_{\text{sk}} \\ m \notin \mathcal{Q}(\mathcal{A}_{\text{sk}}) \end{array} \right]$$

where  $\mathcal{Q}(\mathcal{A}_{\text{sk}})$  denotes the set of queries that  $\mathcal{A}_{\text{sk}}$  makes to the oracle during its strategy.

**NB:** In the case that the secret key should not be known by the verifier, one can use a signature scheme instead.

**Definition 4.1.7** (Signature Scheme). A *signature scheme* SIG is defined by the schema:

$\text{SecretKey} : \text{Type}$   
 $\text{PublicKey} : \text{Type}$   
 $\text{Message} : \text{Type}$   
 $\text{Signature} : \text{Type}$   
 $\text{derive} : \text{SecretKey} \rightarrow \text{PublicKey}$   
 $\text{sign} : \text{SecretKey} \times \text{Message} \rightarrow \mathcal{D}(\text{Signature})$   
 $\text{verify} : \text{PublicKey} \times \text{Message} \times \text{Signature} \rightarrow \text{Bool}$

with the following properties:

- **Completeness:** For all  $\text{sk} : \text{SecretKey}$ ,  $m : \text{Message}$ , and any signature  $\sigma \sim \text{sign}(\text{sk}, m)$ , we have that  $\text{verify}(\text{derive}(\text{sk}), m, \sigma) = \text{True}$ .

**Definition 4.1.8** (Symmetric-Key Encryption Scheme). An *authenticated one-time symmetric-key encryption scheme* **SYM** is defined by the schema:

$$\begin{aligned} &\text{Key} : \text{Type} \\ &\text{Plaintext} : \text{Type} \\ &\text{Ciphertext} : \text{Type} \\ &\text{encrypt} : \text{Key} \times \text{Plaintext} \rightarrow \text{Ciphertext} \\ &\text{decrypt} : \text{Key} \times \text{Ciphertext} \rightarrow \text{Option}(\text{Plaintext}) \end{aligned}$$

with the following properties:

- **Soundness:** For all keys  $k : \text{Key}$  and plaintexts  $p : \text{Plaintext}$ , we have that

$$\text{decrypt}(k, \text{encrypt}(k, p)) = \text{Some}(p)$$

- **Security Requirement:** The symmetric-key encryption scheme must be one-time (INT-CTXT  $\wedge$  IND-CPA)-secure [3]. “One-time” means that an honest protocol participant will almost surely encrypt only one message with a given key; however, the adversary could make many adaptive chosen ciphertext queries for a given key.

**Definition 4.1.9** (Message Digest Cipher). A *message digest cipher* **MDC** is the following scheme:

$$\begin{aligned} &\text{SecretKey} : \text{Type} \\ &\mathbb{F} : \text{Type} \\ &H : \text{SecretKey} \times \mathbb{F} \rightarrow \mathbb{F} \\ &(- \oplus -) : \mathbb{F}^2 \rightarrow \mathbb{F} \\ &(- \ominus -) : \mathbb{F}^2 \rightarrow \mathbb{F} \end{aligned}$$

where  $H$  is a suitable Hash Function and  $\oplus$  and  $\ominus$  fulfill the following right-invertibility constraint for all  $x, y : \mathbb{F}$

$$(x \oplus y) \ominus y = x$$

From these primitives we build a block-based encryption scheme. Given a plaintext  $P : \mathbb{F}^k$  and a key schedule  $K : \text{SecretKey}^k$  we can encrypt to a ciphertext  $C : \mathbb{F}^k$  by sequentially applying the following for each round:

$$C_i := P_i \oplus H(K_i, C_{i-1})$$

where on the first round  $C_{-1}$  is some secret initialization vector. For decryption we use the reverse operation

$$P_i := C_i \ominus H(K_i, C_{i-1})$$

**NB:** This encryption scheme is not *authenticating* in-and-of-itself and requires some strategy like Encrypt-then-MAC for authentication.

**Definition 4.1.10** (Hybrid Public Key Encryption Scheme). A *hybrid public key encryption scheme* [1] **HPKE** is an encryption scheme made up of a symmetric-key encryption scheme **SYM**, a key-agreement scheme **KA**, and a key-derivation function **KDF** to convert from **KA.SharedSecret** to **SYM.Key**. We can define the following encryption and decryption algorithms:

- **Encryption:** Given an ephemeral secret key  $\text{esk} : \text{KA.SecretKey}$ , a public key  $\text{pk} : \text{KA.PublicKey}$ , and plaintext  $p : \text{SYM.Plaintext}$ , we produce the pair

$$m : \text{KA.PublicKey} \times \text{SYM.Ciphertext} := (\text{KA.derive}(\text{esk}), \text{SYM.encrypt}(\text{KDF}(\text{KA.agree}(\text{esk}, \text{pk})), p))$$

- **Decryption:** Given a secret key  $\text{sk} : \text{KA.SecretKey}$ , and an encrypted message, as above,  $m := (\text{epk}, c)$ , we can decrypt  $m$ , producing the plaintext,

$$p : \text{Option}(\text{SYM.Plaintext}) := \text{SYM.decrypt}(\text{KDF}(\text{KA.agree}(\text{sk}, \text{epk})), c)$$

which should decrypt successfully if the **KA.PublicKey** that  $m$  was encrypted with is the derived key of  $\text{sk} : \text{KA.SecretKey}$ .

**Notation:** We denote the above *encrypted message* type as  $\text{Encrypted}(\text{SYM.Plaintext}) := \text{KA.PublicKey} \times \text{SYM.Ciphertext}$ , and the above two algorithms by

$$\begin{aligned} \text{encrypt} &: \text{KA.SecretKey} \times \text{KA.PublicKey} \times \text{SYM.Plaintext} \rightarrow \text{Encrypted}(\text{SYM.Plaintext}) \\ \text{decrypt} &: \text{KA.SecretKey} \times \text{KA.PublicKey} \times \text{SYM.Ciphertext} \rightarrow \text{Option}(\text{SYM.Plaintext}) \end{aligned}$$

**Security Properties:** The HPKE constructed from KA, KDF, and SYM is required to be CCA2-secure and key-private [2].

**Definition 4.1.11** (Authenticated Hybrid Public Key Encryption Scheme). An *authenticated hybrid public encryption scheme* aHPKE is an authenticated encryption scheme built off of an HPKE and a MAC used in the following way:

$$\text{aHPKE.encrypt} : \text{KA.SecretKey} \times \text{KA.PublicKey} \times \text{SYM.Plaintext} \rightarrow \text{AuthEncrypted}(\text{SYM.Plaintext})$$

where AuthEncrypted is the encrypted note type:

$$\text{AuthEncrypted}(\text{SYM.Plaintext}) := \text{MAC.Tag} \times \text{Encrypted}(\text{SYM.Plaintext})$$

and the tag is computed by applying the MAC onto the encrypted note:

$$\text{tag} := \text{MAC}(\text{sk}, \text{HPKE.encrypt}(\text{esk}, \text{pk}))$$

**Definition 4.1.12** (Dynamic Cryptographic Accumulator). A *dynamic cryptographic accumulator* DCA is defined by the schema:

$$\begin{aligned} \text{Item} &: \text{Type} \\ \text{Output} &: \text{Type} \\ \text{Witness} &: \text{Type} \\ \text{State} &: \text{Type} \\ \text{current} &: \text{State} \rightarrow \text{Output} \\ \text{insert} &: \text{Item} \times \text{State} \rightarrow \text{State} \\ \text{prove} &: \text{Item} \times \text{State} \rightarrow \text{Option}(\text{Output} \times \text{Witness}) \\ \text{verify} &: \text{Item} \times \text{Output} \times \text{Witness} \rightarrow \text{Bool} \end{aligned}$$

with the following properties:

- **Unique Accumulated Values:** For any initial state  $s : \text{State}$  and any list of items  $I : \text{List}(\text{Item})$  we can generate the sequence of states:

$$s_0 := s, \quad s_{i+1} := \text{insert}(I_i, s_i)$$

Then, if we collect the accumulated values for these states,  $z_i := \text{current}(s_i)$ , there should be exactly  $|I|$ -many unique values, one for each state update.

- **Provable Membership:** For any initial state  $s : \text{State}$  and any list of items  $I : \text{List}(\text{Item})$  we can generate the sequences of states:

$$s_0 := s, \quad s_{i+1} := \text{insert}(I_i, s_i)$$

Then, if we collect the states  $s_i$  into a set  $S$ , we have the following property for all  $s \in S$  and  $t \in I$ ,

$$\text{Some}(z, w) := \text{prove}(t, s), \quad \text{verify}(t, z, w) = \text{True}$$

**Definition 4.1.13** (Non-Interactive Zero-Knowledge Proving System). A *non-interactive zero-knowledge proving system* NIZK is defined by the schema:

$$\begin{aligned} \text{Statement} &: \text{Type} \\ \text{ProvingKey} &: \text{Type} \\ \text{VerifyingKey} &: \text{Type} \\ \text{PublicInput} &: \text{Type} \\ \text{SecretInput} &: \text{Type} \\ \text{Proof} &: \text{Type} \\ \text{keys} &: \text{Statement} \rightarrow \mathcal{D}(\text{ProvingKey} \times \text{VerifyingKey}) \\ \text{prove} &: \text{Statement} \times \text{ProvingKey} \times \text{PublicInput} \times \text{SecretInput} \rightarrow \mathcal{D}(\text{Option}(\text{Proof})) \\ \text{verify} &: \text{VerifyingKey} \times \text{PublicInput} \times \text{Proof} \rightarrow \text{Bool} \end{aligned}$$

**Notation:** We use the following notation for a NIZK:



- We write the Statement and ProvingKey arguments of `prove` in the superscript and subscript respectively,

$$\text{prove}_{\text{pk}}^P(x, w) := \text{prove}(P, \text{pk}, x, w)$$

- We write the VerifyingKey argument of `verify` in the subscript,

$$\text{verify}_{\text{vk}}(x, \pi) := \text{verify}(\text{vk}, x, \pi)$$

- We say that  $(x, w) : \text{PublicInput} \times \text{SecretInput}$  has the property of being a satisfying input whenever

$$\text{satisfying}_{\text{pk}}^P(x, w) := \exists \pi : \text{Proof}, \text{Some}(\pi) \in \text{prove}_{\text{pk}}^P(x, w)$$

Every NIZK has the following properties for a fixed statement  $P : \text{Statement}$  and keys  $(\text{pk}, \text{vk}) \sim \text{keys}(P)$ :

- **Completeness:** For all  $(x, w) : \text{PublicInput} \times \text{SecretInput}$ , if  $\text{satisfying}_{\text{pk}}^P(x, w) = \text{True}$  with proof witness  $\pi$ , then  $\text{verify}_{\text{vk}}(x, \pi) = \text{True}$ .

- **Knowledge Soundness:** For any polynomial-size adversary  $\mathcal{A}$ ,

$$\mathcal{A} : \text{ProvingKey} \times \text{VerifyingKey} \rightarrow \mathcal{D}(\text{PublicInput} \times \text{Proof})$$

there exists a polynomial-size extractor  $\mathcal{E}_{\mathcal{A}}$

$$\mathcal{E}_{\mathcal{A}} : \text{ProvingKey} \times \text{VerifyingKey} \rightarrow \mathcal{D}(\text{SecretInput})$$

such that the following probability is negligible:

$$\Pr \left[ \begin{array}{l} \text{satisfying}_{\text{pk}}^P(x, w) = \text{False} \\ \text{verify}_{\text{vk}}(x, w) = \text{True} \end{array} \middle| \begin{array}{l} (\text{pk}, \text{vk}) \sim \text{keys}(P) \\ (x, \pi) \sim \mathcal{A}(\text{pk}, \text{vk}) \\ w \sim \mathcal{E}_{\mathcal{A}}(\text{pk}, \text{vk}) \end{array} \right]$$

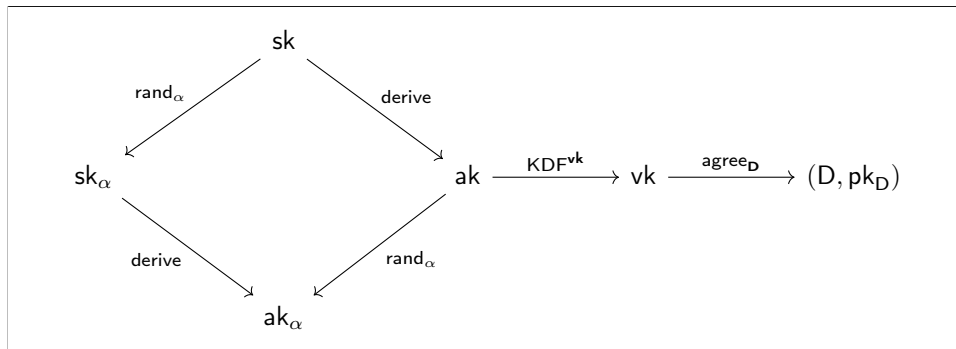
- **Statistical Zero-Knowledge:** There exists a stateful simulator  $\mathcal{S}$ , such that for all stateful distinguishers  $\mathcal{D}$ , the difference between the following two probabilities is negligible:

$$\Pr \left[ \begin{array}{l} \text{satisfying}_{\text{pk}}^P(x, w) = \text{True} \\ \mathcal{D}(\pi) = \text{True} \end{array} \middle| \begin{array}{l} (\text{pk}, \text{vk}) \sim \text{keys}(P) \\ (x, w) \sim \mathcal{D}(\text{pk}, \text{vk}) \\ \text{Some}(\pi) \sim \text{prove}_{\text{pk}}^P(x, w) \end{array} \right] \text{ and } \Pr \left[ \begin{array}{l} \text{satisfying}_{\text{pk}}^P(x, w) = \text{True} \\ \mathcal{D}(\pi) = \text{True} \end{array} \middle| \begin{array}{l} (\text{pk}, \text{vk}) \sim \mathcal{S}(P) \\ (x, w) \sim \mathcal{D}(\text{pk}, \text{vk}) \\ \pi \sim \mathcal{S}(x) \end{array} \right]$$

- **Succinctness:** For all  $(x, w) : \text{PublicInput} \times \text{SecretInput}$ , if  $\text{Some}(\pi) \sim \text{prove}(P, \text{pk}, x, w)$ , then  $|\pi| = \mathcal{O}(1)$ , and  $\text{verify}(\text{vk}, x, \pi)$  runs in time  $\mathcal{O}(|x|)$ .

## 4.2 Addresses and Key Components

For the Transfer protocol we use a multi-layered system of keys:



**Figure 3:** Detailed Key Schedule for MantaPay.

Here we define each key and its function in the Transfer protocol.

**Definition 4.2.1** (Spending Key). Given a key-agreement scheme KA with compatible randomized key-derivation function  $\text{rKDF}$  we have:

$$\text{SpendingKey} := \text{KA.SecretKey}$$

which can be randomized by  $\text{rKDF.Randomness}$ . The spending key must also be compatible with a signature scheme SIG whose underlying key derivation matches KA and  $\text{rKDF}$ .

**Definition 4.2.2** (Proof Authorizing Key). Given a key-agreement scheme  $\text{KA}$  with compatible randomized key-derivation function  $\text{rKDF}$  and with a signature scheme  $\text{SIG}$  we have:

$$\text{ProofAuthorizingKey} := \text{KA.PublicKey}$$

where a given  $\text{sk} : \text{SpendingKey}$  derives the proof-authorizing key by

$$\text{ak} := \text{KA.derive}(\text{sk})$$

This key can be randomized by  $\alpha : \text{rKDF.Randomness}$  to get

$$\text{ak}_\alpha := \text{rKDF.rand}_\alpha(\text{ak})$$

which is also equal to the key-agreement derivation of the randomized secret key

$$\text{ak}_\alpha = \text{KA.derive}(\text{rKDF.rand}_\alpha(\text{sk}))$$

To authorize a message  $m$ , the owner of  $\text{sk}$  can perform the following signature algorithm:

$$\sigma \sim \text{SIG.sign}(\text{rKDF.rand}_\alpha(\text{sk}), m)$$

which can then be verified against  $\text{ak}_\alpha$  with

$$\text{SIG.verify}(\text{ak}_\alpha, m, \sigma)$$

**NB:** For the **Transfer** protocol, the message will be the zero-knowledge proof of a valid transfer and any additional associated data and **Ledger** payload.

**Definition 4.2.3** (Viewing Key). Given a proof-authorizing key, we require a  $\text{KDF}^{\text{vk}}$  of type

$$\text{KDF}^{\text{vk}} : \text{ProofAuthorizingKey} \rightarrow \text{ViewingKey}$$

where  $\text{ViewingKey}$  is of type  $\text{KA.SecretKey}$  so that it can be available for the  $\text{KA}$  key-agreement scheme so that we have

$$\text{vk} := \text{KDF}^{\text{vk}}(\text{ak})$$

**Definition 4.2.4** (Shielded Address). Given a viewing key which is the secret key for the key-agreement scheme  $\text{KA}$ , the shielded address is given by randomly selecting a public key  $D : \text{KA.PublicKey}$  and then performing  $\text{KA.agree}$  against it:

$$\text{pk}_D := \text{KA.agree}_D(\text{vk})$$

We return the pair as the shielded address:  $\text{addr} := (D, \text{pk}_D)$ . We call the random element  $D$  the **Diversifier** for the shielded address  $\text{addr}$ .

**Definition 4.2.5** (Key Schedule). A **KeySchedule** is a collection of implementations of the following abstract cryptographic primitives as described in the above definitions:

- **Key-Agreement Scheme:**  $\text{KA}$
- **Randomized Key-Derivation Scheme:**  $\text{rKDF}$
- **Viewing Key Derivation Function:**  $\text{KDF}^{\text{vk}}$
- **Proof Authorization Signature:**  $\text{SIG}$

with the following notational conventions:

$$\begin{aligned} \text{SpendingKey} &:= \text{KA.SecretKey} \\ \text{ProofAuthorizingKey} &:= \text{KA.PublicKey} \\ \text{ViewingKey} &:= \text{KA.SecretKey} \\ \text{Diversifier} &:= \text{KA.PublicKey} \\ \text{ShieldedAddress} &:= \text{KA.Diversifier} \times \text{KA.PublicKey} \end{aligned}$$

with the following constraints:

$$\begin{aligned}
\text{rKDF.Input} &= \text{KA.SecretKey} \\
\text{rKDF.Output} &= \text{KA.PublicKey} \\
\text{SIG.SecretKey} &= \text{KA.SecretKey} \\
\text{SIG.PublicKey} &= \text{KA.PublicKey} \\
\text{SIG.derive} &= \text{KA.derive}
\end{aligned}$$

### 4.3 Transfer Protocol

The **Transfer** protocol is the fundamental abstraction in **MantaPay** and facilitates the valid transfer of **Assets** among participants while preserving their privacy. The **Transfer** is made up of sub-components called **Senders** and **Receivers** which represent the private input and the private output of a transaction.<sup>4</sup> To perform a **Transfer**, a protocol participant gathers the **SpendingKeys** they own, selects a subset of the **UTXOs** they have still not spent (with a fixed **AssetId**), collects **ShieldedAddresses** from other participants for the outputs of the **Transfer**, assigning each key a subset of the input **Assets**, and then builds a **Transfer** object representing that transaction. From this **Transfer** object, they construct a **TransferPost** which they then send to the **Ledger** to be validated, representing a completed state transition in the **Ledger**, updating the **UTXOSet** and **VoidNumberSet**. The transformation from **Transfer** to **TransferPost** involves keeping the parts of the **Transfer** that *must* be known to the **Ledger** and for the parts that *should not* be known, substituting them for a *zero-knowledge proof* representing the validity of the secret information known to the participant, and the **Transfer** as a whole.

We begin by defining the cryptographic primitives involved in the **Transfer** protocol:

**Definition 4.3.1** (Transfer Configuration). A **TransferConfiguration** is a collection of implementations of the following abstract cryptographic primitives:

- **Key Schedule:** **KeySchedule**
- **Incoming Authenticated Hybrid Public Key Encryption:**  $\text{aHPKE}^{\text{in}}$
- **Outgoing Authenticated Hybrid Public Key Encryption:**  $\text{aHPKE}^{\text{out}}$
- **UTXO Commitment Scheme:**  $\text{COM}^{\text{UTXO}}$
- **Void Number Commitment Scheme:**  $\text{COM}^{\text{VN}}$
- **Dynamic Cryptographic Accumulator:** **DCA**
- **Zero-Knowledge Proving System:** **NIZK**

with the following notational conventions:

$$\begin{aligned}
\text{UTXO} &:= \text{COM}^{\text{UTXO}}.\text{Output} \\
\text{VoidNumber} &:= \text{COM}^{\text{VN}}.\text{Output} \\
\text{IncomingNote} &:= \text{KeySchedule.ShieldedAddress} \times \text{COM}^{\text{UTXO}}.\text{Randomness} \times \text{Asset} \\
\text{OutgoingNote} &:= \text{Asset} \\
\text{UTXOSet} &:= \text{DCA}
\end{aligned}$$

and the following constraints:

$$\begin{aligned}
\text{COM}^{\text{UTXO}}.\text{Input} &= \text{KeySchedule.ShieldedAddress} \times \text{Asset} \\
\text{COM}^{\text{VN}}.\text{Randomness} &= \text{KeySchedule.ProofAuthorizingKey} \\
\text{COM}^{\text{VN}}.\text{Input} &= \text{UTXO} \\
\text{UTXOSet.Item} &= \text{UTXO} \\
\text{aHPKE}^{\text{in}}.\text{KA} &= \text{KeySchedule.KA} \\
\text{aHPKE}^{\text{out}}.\text{KA} &= \text{KeySchedule.KA} \\
\text{ValidTransfer} &: \text{NIZK.Statement}
\end{aligned}$$

where **ValidTransfer** is defined below.

<sup>4</sup>Note that they do not represent actual individual participants in a transaction, but instead just the data involved in the transaction.

For the rest of this section, we assume the existence of a `TransferConfiguration` and use the primitives outlined above explicitly. We also implicitly use the `KeySchedule` and drop its prefix when referring to its members. We continue by defining the `Sender` and `Receiver` constructions as well as their public counterparts, the `SenderPost` and `ReceiverPost`.

**Definition 4.3.2** (Transfer Sender). A `Sender` is the following tuple:

$$\begin{aligned} & \text{ak} : \text{ProofAuthorizingKey} \\ & \alpha : \text{rKDF.Randomness} \\ & \text{ak}_\alpha : \text{KA.PublicKey} \\ & D : \text{Diversifier} \\ & r : \text{COM}^{\text{UTXO}}.\text{Randomness} \\ & \text{asset} : \text{Asset} \\ & \text{pk}_D : \text{KA.PublicKey} \\ & \text{esk}_{\text{out}} : \text{KA.SecretKey} \\ & (\text{tag}_{\text{out}}, \text{epk}_{\text{out}}, C_{\text{out}}) : \text{AuthEncrypted}(\text{OutgoingNote}) \\ & \text{cm} : \text{UTXO} \\ & (z_{\text{cm}}, \pi_{\text{cm}}) : \text{UTXOSet.MembershipProof} \\ & \text{vn} : \text{VoidNumber} \end{aligned}$$

A `Sender`,  $S$ , is constructed from a proof authorizing key  $\text{ak} : \text{ProofAuthorizingKey}$ , a randomizer  $\alpha : \text{rKDF.Randomness}$  and an encrypted message  $(\text{tag}_{\text{in}}, \text{epk}_{\text{in}}, C_{\text{in}}) : \text{AuthEncrypted}(\text{IncomingNote})$  with the following algorithm:

$$\begin{aligned} & \text{ak}_\alpha := \text{rKDF.rand}_\alpha(\text{ak}) \\ & \text{vk} := \text{KDF}^{\text{vk}}(\text{ak}) \\ & \text{Some}(D, r, \text{asset}) := \text{aHPKE}^{\text{in}}.\text{decrypt}(\text{vk}, \text{tag}_{\text{in}}, \text{epk}_{\text{in}}, C_{\text{in}}) \\ & \text{pk}_D := \text{KA.agree}_D(\text{vk}) \\ & \text{esk}_{\text{out}} : \sim \text{KA.SecretKeyDistribution} \\ & (\text{tag}_{\text{out}}, \text{epk}_{\text{out}}, C_{\text{out}}) := \text{aHPKE}^{\text{out}}.\text{encrypt}_D(\text{esk}_{\text{out}}, \text{pk}_D, \text{asset}) \\ & \text{cm} := \text{COM}_r^{\text{UTXO}}(D, \text{pk}_D, \text{asset}) \\ & \text{Some}(z_{\text{cm}}, \pi_{\text{cm}}) := \text{UTXOSet.prove}(\text{cm}) \\ & \text{vn} := \text{COM}_{\text{ak}}^{\text{VN}}(\text{cm}) \end{aligned}$$

**Definition 4.3.3** (Transfer Sender Post). A `SenderPost` is the following tuple extracted from a `Sender`:

$$\begin{aligned} & \text{ak}_\alpha : \text{KA.PublicKey} \\ & (\text{tag}_{\text{out}}, \text{epk}_{\text{out}}, C_{\text{out}}) : \text{AuthEncrypted}(\text{OutgoingNote}) \\ & z_{\text{cm}} : \text{UTXOSet.Output} \\ & \text{vn} : \text{VoidNumber} \end{aligned}$$

which are the parts of a `Sender` which should be *posted* to the `Ledger`.

**Definition 4.3.4** (Transfer Receiver). A `Receiver` is the following tuple:

$$\begin{aligned} & (D, \text{pk}_D) : \text{ShieldedAddress} \\ & r : \text{COM}^{\text{UTXO}}.\text{Randomness} \\ & \text{asset} : \text{Asset} \\ & \text{cm} : \text{UTXO} \\ & \text{esk}_{\text{in}} : \text{KA.SecretKey} \\ & (\text{tag}_{\text{in}}, \text{epk}_{\text{in}}, C_{\text{in}}) : \text{AuthEncrypted}(\text{IncomingNote}) \end{aligned}$$

A `Receiver`,  $R$ , is constructed from a shielded address  $(D, \text{pk}_D) : \text{ShieldedAddress}$ , an asset  $\text{asset} : \text{Asset}$ , and a

UTXO-commitment randomness  $r : \text{COM}^{\text{UTXO}}$ . Randomness with the following algorithm:

$$\begin{aligned} \text{cm} &:= \text{COM}_r^{\text{UTXO}}(\text{D}, \text{pk}_\text{D}, \text{asset}) \\ \text{esk}_\text{in} &:= \text{KA.SecretKeyDistribution} \\ (\text{tag}_\text{in}, \text{epk}_\text{in}, \text{C}_\text{in}) &:= \text{aHPKE}^\text{in}.\text{encrypt}_\text{D}(\text{esk}_\text{in}, \text{pk}_\text{D}, (\text{D}, r, \text{asset})) \end{aligned}$$

**Definition 4.3.5** (Transfer Receiver Post). A ReceiverPost is the following tuple extracted from a Receiver:

$$\begin{aligned} \text{cm} &: \text{UTXO} \\ (\text{tag}_\text{in}, \text{epk}_\text{in}, \text{C}_\text{in}) &: \text{AuthEncrypted}(\text{IncomingNote}) \end{aligned}$$

which are the parts of a Receiver which should be *posted* to the Ledger.

**Definition 4.3.6** (Transfer Sources and Sinks). A Source (or a Sink) is an Asset representing a public input (or output) of a Transfer.

**Definition 4.3.7** (Transfer Object). A Transfer is the following tuple:

$$\begin{aligned} \text{sources} &: \text{List}(\text{Asset}) \\ \text{senders} &: \text{List}(\text{Sender}) \\ \text{receivers} &: \text{List}(\text{Receiver}) \\ \text{sinks} &: \text{List}(\text{Asset}) \end{aligned}$$

The *shape* of a Transfer is the following 4-tuple of cardinalities of those sets

$$(|T.\text{sources}|, |T.\text{senders}|, |T.\text{receivers}|, |T.\text{sinks}|)$$

In order for a Transfer to be considered *valid*, it must adhere to the following constraints:

- **Same Id**: All the AssetIds in the Transfer must be equal.
- **Balanced**: The sum of input AssetValues must be equal to the sum of output AssetValues.
- **Well-formed Senders**: All of the Senders in the Transfer must be constructed according to the above Sender definition.
- **Well-formed Receivers**: All of the Receivers in the Transfer must be constructed according to the above Receiver definition.

In order to prove that these constraints are satisfied for a given Transfer, we build a zero-knowledge proof which will witness that the Transfer is valid and should be accepted by the Ledger.

**Definition 4.3.8** (Transfer Validity Statement). A transfer  $T : \text{Transfer}$  is considered *valid* if and only if

1. All the AssetIds in  $T$  are equal:

$$\left| \left( \bigcup_{a \in T.\text{sources}} a.\text{id} \right) \cup \left( \bigcup_{S \in T.\text{senders}} S.\text{asset}.\text{id} \right) \cup \left( \bigcup_{R \in T.\text{receivers}} R.\text{asset}.\text{id} \right) \cup \left( \bigcup_{a \in T.\text{sinks}} a.\text{id} \right) \right| = 1$$

2. The sum of input AssetValues is equal to the sum of output AssetValues:

$$\left( \sum_{a \in T.\text{sources}} a.\text{value} \right) + \left( \sum_{S \in T.\text{senders}} S.\text{asset}.\text{value} \right) = \left( \sum_{R \in T.\text{receivers}} R.\text{asset}.\text{value} \right) + \left( \sum_{a \in T.\text{sinks}} a.\text{value} \right)$$

3. For all  $S \in T.\text{senders}$ , the Sender  $S$  is well-formed:

$$\begin{aligned} \text{ak}_\alpha &= \text{rKDF}.\text{rand}_\alpha(\text{ak}) \\ \text{vk} &= \text{KDF}^{\text{vk}}(\text{ak}) \\ \text{pk}_\text{D} &= \text{KA}.\text{derive}_\text{D}(\text{vk}) \\ (\text{tag}_\text{out}, \text{epk}_\text{out}, \text{C}_\text{out}) &= \text{aHPKE}^\text{out}.\text{encrypt}_\text{D}(\text{esk}_\text{out}, \text{pk}_\text{D}, \text{asset}) \\ \text{cm} &= \text{COM}_r^{\text{UTXO}}(\text{D}, \text{pk}_\text{D}, \text{asset}) \\ \text{UTXOSet}.\text{verify}(\text{cm}, z_\text{cm}, \pi_\text{cm}) &= \text{True} \\ \text{vn} &= \text{COM}_{\text{ak}}^{\text{VN}}(\text{cm}) \end{aligned}$$

4. For all  $R \in T.\text{receivers}$ , the Receiver  $R$  is well-formed:

$$\begin{aligned} \text{cm} &= \text{COM}_r^{\text{UTXO}}(\text{D}, \text{pk}_\text{D}, \text{asset}) \\ (\text{tag}_\text{in}, \text{epk}_\text{in}, \text{C}_\text{in}) &= \text{aHPKE}^\text{in}.\text{encrypt}_\text{D}(\text{esk}_\text{in}, (\text{D}, r, \text{asset})) \end{aligned}$$

**Notation:** This statement is denoted `ValidTransfer` and is assumed to be expressible as a Statement of NIZK.

**Definition 4.3.9** (Transfer Post). A `TransferPost` is the following tuple:

$$\begin{aligned} \text{sources} &: \text{List}(\text{Source}) \\ \text{senders} &: \text{List}(\text{SenderPost}) \\ \text{receivers} &: \text{List}(\text{ReceiverPost}) \\ \text{sinks} &: \text{List}(\text{Sink}) \\ \pi &: \text{NIZK.Proof} \end{aligned}$$

A `TransferPost`,  $P$ , is constructed by assembling the zero-knowledge proof of `Transfer` validity from a known proving key  $\text{pk} : \text{NIZK.ProvingKey}$  and a given  $T : \text{Transfer}$ :

$$\begin{aligned} x &:= \text{Transfer.public}(T) \\ w &:= \text{Transfer.secret}(T) \\ \text{Some}(\pi) &\sim \text{NIZK.prove}_{\text{pk}}^{\text{ValidTransfer}}(x, w) \\ P.\text{sources} &:= x.\text{sources} \\ P.\text{senders} &:= x.\text{senders} \\ P.\text{receivers} &:= x.\text{receivers} \\ P.\text{sinks} &:= x.\text{sinks} \\ P.\pi &:= \pi \end{aligned}$$

where `Transfer.public` returns `SenderPosts` for each `Sender` in  $T$  and `ReceiverPosts` for each `Receiver` in  $T$ , keeping `Sources` and `Sinks` as they are, and `Transfer.secret` returns all the rest of  $T$  which is not part of the output of `Transfer.public`.

Now that the prover has constructed the proof, the underlying spending keys need to authorize the transaction before it can be sent to the `Ledger`.

**Definition 4.3.10** (Proof Authorization). Given a transfer post  $T : \text{TransferPost}$  and a set of spending keys  $S = \{(\text{sk}_i, \alpha_i)\}$  where  $(\text{sk}_i, \alpha_i)$  come from the  $i$ th spender associated to the  $T.\text{senders}_i$ , we have the following signature:

$$\Sigma := \{\text{SIG.sign}(\text{rKDF.rand}_\alpha(\text{sk}), T) \mid (\text{sk}, \alpha) \in S\}$$

which can be verified by the ledger with

$$\forall_i \text{SIG.verify}(T.\text{senders}_i.\text{ak}_\alpha, T, \Sigma_i) = \text{True}$$

Now that the transfer post has been signed by the owners of the spending keys, it can be sent up to the `Ledger`.

**Definition 4.3.11** (Ledger-side Transfer Validity). To check that  $P$  represents a valid `Transfer`, the ledger checks the following:

- **Signature Check:** All the signatures associated to the transactions are valid.
- **Public Withdraw:** All the public addresses corresponding to the `Assets` in  $P.\text{sources}$  have enough public balance (i.e. in the `PublicLedger`) to withdraw the given `Asset`.
- **Public Deposit:** All the public addresses corresponding to the `Assets` in  $P.\text{sinks}$  exist.
- **Current Accumulated State:** The `UTXOSet.Output` stored in each  $P.\text{senders}$  is equal to current accumulated value, `UTXOSet.current(Ledger.utxos())`, for the current state of the `Ledger`.
- **New VoidNumbers:** All the `VoidNumbers` in  $P.\text{senders}$  are unique, and no `VoidNumber` in  $P.\text{senders}$  has already been stored in the `Ledger.VoidNumberSet`.
- **New UTXOs:** All the `UTXOs` in  $P.\text{receivers}$  are unique, and no `UTXO` in  $P.\text{receivers}$  has already been stored on the ledger.

- **Verify Transfer:** Check that  $\text{NIZK.verify}_{\text{vk}}(P.\text{sources} \parallel P.\text{senders} \parallel P.\text{receivers} \parallel P.\text{sinks}, P.\pi) = \text{True}$ .

**Definition 4.3.12** (Ledger Transfer Update). After checking that a given *TransferPost*  $P$  is valid, the *Ledger* updates its state by performing the following changes:

- **Public Updates:** All the relevant public accounts on the *PublicLedger* are updated to reflect their new balances using the *Sources* and *Sinks* present in  $P$ .
- **UTXOSet Update:** The new UTXOs are appended to the *UTXOSet*.
- **VoidNumberSet Update:** The new *VoidNumbers* are appended to the *VoidNumberSet*.

## 4.4 Batched Transactions

For *MantaPay* participants to use the *Transfer* protocol, they will need to keep track of the current state of their shielded assets and use them to build *TransferPosts* to send to the *Ledger*. The *shielded balance* of any participant is the sum of the balances of their shielded assets, but this balance may be fragmented into arbitrarily many pieces, as each piece represents an independent asset that the participant received as the output of some *Transfer*. To then spend a subset of their shielded balance, the participant would need to accumulate all of the relevant fragments into a large enough *shielded asset* to spend all at once, building a collection of *TransferPosts* to send to the *Ledger*.

---

### Algorithm 1 Batch Transaction Algorithm

---

```

procedure BUILDTRANSACTION(sk,  $\mathcal{B}$ , total, addr)
   $B \leftarrow \text{Sample}(\text{total}, \mathcal{B})$  ▷ Samples key-asset pairs from  $\mathcal{B}$  whose asset total at least total
  if  $\text{len}(B) = 0$  then
    return [] ▷ Insufficient Balance
  end if
   $P \leftarrow []$  ▷ Allocate a new list for TransferPosts
  while  $\text{len}(B) > N$  do ▷ While there are enough pairs to make another Transfer
     $A \leftarrow []$ 
    for  $b \in (B, N)$  do ▷ Get the next  $N$  pairs from  $B$ 
       $S \leftarrow \text{BuildSenders}_{\text{sk}}(b)$ 
       $[acc, zs...] \leftarrow \text{BuildAccumulatorAndZeroes}_{\text{sk}}(S)$  ▷ Build a new accumulator and zeroes
       $P \leftarrow P + \text{TransferPost}(\text{Transfer}([], S, [acc, zs...], []))$ 
       $(A, Z) \leftarrow (A + (acc.\tilde{d}, acc.\text{asset.value}), Z + zs)$  ▷ Save  $acc$  for the next loop,  $zs$  for the end
    end for
     $B \leftarrow A + \text{remainder}(B, N)$ 
  end while
   $S \leftarrow \text{PrepareZeroes}_{\text{sk}}(N, B, Z, P)$  ▷ Use  $Z$  and Mints to make  $B$  go up to  $N$  in size.
   $R \leftarrow \text{BuildReceiver}_{\text{sk}}(\text{addr}, S)$ 
   $[c, zs...] \leftarrow \text{BuildAccumulatorAndZeroes}_{\text{sk}}(S)$ 
  return  $P + \text{TransferPost}(\text{Transfer}([], S, [R, c, zs...], []))$ 
end procedure

```

---

Any wallet implementation should see that their users need not keep track of this complexity themselves. Instead, like a public ledger, the notion of a *transaction* between one participant and another should be viewed as a single (atomic) action that the user can take, performing a withdrawal from their shielded balance. To describe such a *semantic transaction*, we assume the existence of two transfer shapes<sup>5</sup>: *Mint* with shape  $(1, 0, 1, 0)$  and *PrivateTransfer* with shape  $(0, N, N, 0)$  for some natural number  $N > 1$ .

For a fixed spending key,  $\text{sk} : \text{SpendingKey}$ , and asset id,  $\text{id} : \text{AssetId}$ , we are given a balance state,  $\mathcal{B} : \text{FinSet}(\text{KA.PublicKey} \times \text{AssetValue})$ , a set of key-asset pairs for unspent assets, a total balance to withdraw,  $\text{total} : \text{AssetValue}$ , and a shielded key  $\text{addr} : \text{ShieldedAddress}$ . We can then compute

$\text{BUILDTRANSACTION}(\text{sk}, \mathcal{B}, \text{total}, \text{addr})$

to receive a  $\text{List}(\text{TransferPost})$  to send to the ledger, representing the transfer of  $\text{total}$  to  $\text{addr}$ .

If all of the *Transfers* are accepted by the ledger, the balance state  $\mathcal{B}$  should be updated accordingly, removing all of the pairs which were used in the *Transfer*. Wallets should also handle the more complex case when only some of the *Transfers* succeed in which case they need to be able to continue retrying the transaction until they

---

<sup>5</sup>Other *Transfer* accumulation algorithms are possible with different starting shapes.

are finally resolved. Since the only **Transfer** which sends **Assets** out of the control of the user is the last one (and it recursively depends on the previous **Transfers**), then it is safe to continue from a partially resolved state with a simple retry of the **BUILDTRANSACTION** algorithm.

## 5 Concrete Protocol

We define the instantiation of the abstract protocol in this section, but first some preliminary notes.

### 5.1 Poseidon Permutation

The **Poseidon** Permutation [5] is a finite field cryptographic primitive that can be used in lots of different contexts, like hash functions, commitment schemes, and symmetric encryption. **Poseidon** plays a fundamental role in simplifying the **Transfer** protocol and reducing the overall cost of the Zero-Knowledge circuits. **Poseidon** (without sponges) is a family of hash functions with the following signature:

$$\mathbf{Poseidon}_k : \mathbb{F} \times \mathbb{F}^k \rightarrow \mathbb{F}$$

over some sufficiently large finite field  $\mathbb{F}$ . The first distinguished field element is used as a domain separation element. For this purpose, we use the following hashing function to generate domain strings:

$$\text{HashToScalar}(m) := \mathbb{F}.\text{truncate}(\text{Blake2s}(m))$$

We make use of **Poseidon** for a few values of  $k$  in the concrete protocol below.

### 5.2 Elliptic Curve Cryptography

Because we use a Zero-Knowledge Proving System, we want the cryptographic constructions that feature in our protocol to be *ZKP-friendly*. For a ZKP system defined over a field  $\mathbb{F}$  we can look for elliptic curves that have a base field of the same order as  $\mathbb{F}$ . These such curves are said to be “embeddable” or “embedded in”  $\mathbb{F}$ . For the constructions below, we use  $\mathbb{F}$  as the proof system field and  $\mathbb{G}$  as an embedded curve with scalar field  $\mathbb{S}$ . We also assume that  $|\mathbb{S}| < |\mathbb{F}|$  so we can use the injection lift  $\mathbb{S} \rightarrow \mathbb{F}$  to lift scalars to the proof system field.

To use group elements in affine form we also define the projections:

$$x : \mathbb{G} \rightarrow \mathbb{F} \text{ and } y : \mathbb{G} \rightarrow \mathbb{F}$$

which we use below to insert group elements into field-only hash functions.

For this protocol, we use **BLS12-381** as our outer (pairing-friendly) curve with scalar field  $\mathbb{F}$  and **JubJub** as our inner curve with scalar field  $\mathbb{S}$ .

### 5.3 Concrete Cryptographic Schemes

**Definition 5.3.1** (Commitment Schemes). The protocol features two different commitment schemes:  $\text{COM}^{\text{UTXO}}$  the UTXO Commitment Scheme and  $\text{COM}^{\text{VN}}$  the Void Number Commitment Scheme. Both commitment schemes use **Poseidon** as the underlying cryptographic primitive. The UTXO uses an arity-8 **Poseidon** with the following mapping:

$$\text{COM}_r^{\text{UTXO}}(D, \text{pk}_D, \text{asset}) := \mathbf{Poseidon}_8(d, 0, r, x(D), y(D), x(\text{pk}_D), y(\text{pk}_D), \text{asset.id}, \text{asset.value})$$

where  $d = \text{HashToScalar}(\text{“manta-pay/1.0.0/com-utxo”})$ . For the Void Number Commitment Scheme we use an arity-4 **Poseidon** with the following mapping:

$$\text{COM}_{\text{ak}}^{\text{VN}}(\text{cm}) := \mathbf{Poseidon}_4(\text{HashToScalar}(\text{“manta-pay/1.0.0/com-vn”}), 0, x(\text{ak}), y(\text{ak}), \text{cm})$$

**Definition 5.3.2** (Key-Agreement Scheme). For KA, we use a Diffie-Hellman Key Exchange over  $(\mathbb{G}, \mathbb{S})$ :

$$\begin{aligned} \text{KA.derive}(x) : \mathbb{S} \rightarrow \mathbb{G} &:= x \cdot G \\ \text{KA.agree}(x, Y) : \mathbb{S} \times \mathbb{G} \rightarrow \mathbb{G} &:= x \cdot Y \end{aligned}$$

where  $G$  is a fixed public point.



**Definition 5.3.3** (Symmetric-Key Encryption Scheme). For **SYM** we use **Poseidon**<sub>2</sub> as the hash function in a message digest cipher with key-schedule given by the following:

$$K_i := \text{Poseidon}_2(\text{HashToScalar}(\text{"manta-pay/1.0.0/mdc-key-schedule"}), K_0, K_{i-1})$$

**Definition 5.3.4** (Key-Derivation Functions). For the encryption scheme KDFs, we use the following which maps a group element  $G : \mathbb{G}$  to a scalar:

$$\text{KDF}(G) := \text{Poseidon}_2(\text{HashToScalar}(\text{"manta-pay/1.0.0/encryption-kdf"}), x(G), y(G))$$

**Definition 5.3.5** (Message Authentication Code). For message authentication codes we use the following instantiation of **Poseidon**:

$$\text{MAC}(\text{sk}, m) := \text{Poseidon}_{|m|+1}(\text{HashToScalar}(\text{"manta-pay/1.0.0/mac"}), \text{sk}, m)$$

In this protocol, we use  $|m| \in \{2, 6\}$  for **OutgoingNote** and **IncomingNote** respectively.

**Definition 5.3.6** (Signature Scheme). For the signature scheme we use standard ECDSA over  $\mathbb{G}$ .

**Definition 5.3.7** (Dynamic Cryptographic Accumulator). For DCA, we use a Merkle Tree with **Poseidon**<sub>2</sub> as the inner node combining hash function and no leaf hash function. It is safe to omit the leaf hash function in this case because the leaf values are already the outputs of a hash function and cannot be directly controlled.

**Definition 5.3.8** (Non-Interactive Zero-Knowledge Proving System). For NIZK, the protocol can use any non-interactive zero-knowledge proving system like Groth16 [5] and/or PLONK/PLONKUP [4, 6].

## 5.4 AssetValue Bounds Check

In order to implement the balanced transfer relation one needs to ensure that the amount of input value is equal to the amount of output value. However, since we're working over finite fields, the naïve arithmetic wraps past zero and is vulnerable to range-based attacks. Instead we constrain every **AssetValue** to be less than some bound  $\mathcal{V}$  and that every sum over those values is also less than  $\mathcal{V}$ . Since we're using BLS12-381 we are safe to use  $\mathcal{V} = 2^{128}$ .

## 6 Acknowledgements

We would like to thank Luke Pearson and Toghrul Maharrov for our insightful discussions on reusable shielded addresses.

## 7 References

### References

- [1] Richard Barnes, Karthikeyan Bhargavan, Benjamin Lipp, and Christopher A. Wood. Hybrid Public Key Encryption. Internet-Draft draft-irtf-cfrg-hpke-12, Internet Engineering Task Force, September 2021. Work in Progress.
- [2] Mihir Bellare, Alexandra Boldyreva, Anand Desai, and David Pointcheval. Key-privacy in public-key encryption. In *ASIACRYPT*, volume 2248 of *Lecture Notes in Computer Science*, pages 566–582. Springer, 2001.
- [3] Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In *ASIACRYPT*, volume 1976 of *Lecture Notes in Computer Science*, pages 531–545. Springer, 2000.
- [4] Ariel Gabizon, Zachary J. Williamson, and Oana Ciobotaru. PLONK: permutations over lagrange-bases for oecumenical noninteractive arguments of knowledge. *IACR Cryptol. ePrint Arch.*, page 953, 2019.
- [5] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A new hash function for zero-knowledge proof systems. In *USENIX Security Symposium*, pages 519–535. USENIX Association, 2021.
- [6] Luke Pearson, Joshua Fitzgerald, Héctor Masip, Marta Bellés-Muñoz, and Jose Luis Muñoz-Tapia. Plonkup: Reconciling plonk with plonkup. *IACR Cryptol. ePrint Arch.*, page 86, 2022.