MantaPay Protocol Specification
v0.4.0

Shumo Chu and Brandon H. Gomes

March 21, 2022

Abstract

MantaPay is an implementation of a *decentralized anonymous payment* scheme based on the MANTA$_{\text{DAP}}$ protocol outlined in the original MANTA whitepaper.

Contents

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	✗
Monero	RingCT/NIZK	PoW	1	✗
Tornado Cash (Nova)	NIZK	PoW	2	✓
MantaPay 0.4.0	NIZK	PoS	1	✓

Table 1: Comparison of MantaPay with previous constructions

2 Notation

The following notation is used throughout this specification:

- Type is the type of types[1].
- If $x : T$ then x is a value and T is a type, denoted $T : \mathsf{Type}$, and we say that x *has type* T.
- Bool is the type of booleans with values True and False.
- For any types $A : \mathsf{Type}$ and $B : \mathsf{Type}$ we denote the *type of functions* from A to B as $A \to B : \mathsf{Type}$.
- For any types $A : \mathsf{Type}$ and $B : \mathsf{Type}$ we denote the *product type* over A and B as $A \times B : \mathsf{Type}$ with constructor $(-,-) : A \to (B \to A \times B)$.
- For any type $T : \mathsf{Type}$, we define $\mathsf{Option}(T) : \mathsf{Type}$ as the inductive type with constructors:

$$\mathsf{None} : \mathsf{Option}(T)$$
$$\mathsf{Some} : T \to \mathsf{Option}(T)$$

- We denote the *type of finite sets* over a type $T : \mathsf{Type}$ as $\mathsf{FinSet}(T) : \mathsf{Type}$. The membership predicate for a value $x : T$ in a finite set $S : \mathsf{FinSet}(T)$ is denoted $x \in S$.
- We denote the *type of finite ordered sets* over a type $T : \mathsf{Type}$ as $\mathsf{List}(T) : \mathsf{Type}$. This can either be defined by an inductive type or as a $\mathsf{FinSet}(T)$ with a fixed ordering. We denote the constructor for a list as $[\ldots]$ for an arbitrary set of elements.
- We denote the *type of distributions* over a type $T : \mathsf{Type}$ as $\mathfrak{D}(T) : \mathsf{Type}$. 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, $y \in \{x \,|\, x \sim \mathfrak{D}(T)\} \leftrightarrow y \in \mathfrak{D}(T)$.
- 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 $a : \mathsf{Asset}$ is a tuple

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

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 facilitiates 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*.

[1]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.

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[2]. 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 an *address* which they use as a unique identifier to represent them on the ledger.

$$\begin{tikzcd}
 & \mathsf{sk} \arrow[r, Rightarrow] \arrow[dl, bend right] & \mathsf{pk} \arrow[dr, bend left] & \\
(\mathsf{sk}, \mathsf{svk}) & & & (\mathsf{pk}, \mathsf{pvk}) \\
 & \mathsf{svk} \arrow[r, Rightarrow] \arrow[ul, bend left] & \mathsf{pvk} \arrow[ur, bend right] &
\end{tikzcd}$$

Figure 1: Key pairs and Addresses.

MantaPay uses two kinds of cryptographic keypairs to build an address, *spending keys*, sk and pk, and *viewing keys*, svk and pvk. An address is the pair $(\mathsf{pk}, \mathsf{pvk})$ of public keys. The keys have the following properties:

- Access to a public spending key pk and public viewing key pvk represents the ability to send Assets to the owner of the associated sk.
- Access to a secret viewing key svk represents the ability to reveal shielded Asset information for Assets belonging to the owner of the associated sk.
- Access to a secret spending key sk represents the ability to spend Assets that were received under the associated public spending key pk.

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

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 transfered, but the *ability to spend the asset in the future*, which should be transfered. To enforce this second invariant we can use a public ledger[3] that keeps track of the movement of Assets from one participant 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.

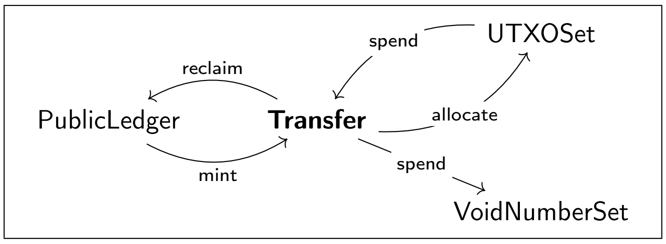

Figure 2: Lifecycle of an Asset.

[2]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*.

[3]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.

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 refering to an allocated Asset transfered 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, 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 $(\mathsf{UTXO}, \mathsf{EncryptedNote})$ pair, the ShieldedAssetPool requires another commitment in order to spend the Asset, transfering 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.

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

$$\begin{aligned} \mathsf{Trapdoor} &: \mathsf{Type} \\ \mathsf{Input} &: \mathsf{Type} \\ \mathsf{Output} &: \mathsf{Type} \\ \mathsf{TrapdoorDistribution} &: \mathfrak{D}(\mathsf{Trapdoor}) \\ \mathsf{commit} &: \mathsf{Trapdoor} \times \mathsf{Input} \to \mathsf{Output} \end{aligned}$$

with the following properties:

- **Binding**: It is infeasible to find an $x, y : \mathsf{Input}$ and $r, s : \mathsf{Trapdoor}$ such that $x \neq y$ and $\mathsf{commit}(r, x) = \mathsf{commit}(s, y)$.

- **Hiding**: For all $x, y : \mathsf{Input}$, the distributions $\{\mathsf{commit}(r, x) \mid r \sim \mathsf{TrapdoorDistribution}\}$ and $\{\mathsf{commit}(r, y) \mid r \sim \mathsf{TrapdoorDistribution}\}$ are *computationally indistinguishable.*

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

Definition 4.1.2 (Hash Function)**.** A *hash function* HASH is defined by the schema:

$$\begin{aligned} \mathsf{Input} &: \mathsf{Type} \\ \mathsf{Output} &: \mathsf{Type} \\ \mathsf{hash} &: \mathsf{Input} \to \mathsf{Output} \end{aligned}$$

with the following properties:

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

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 two parts.

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

Definition 4.1.3 (Key-Agreement Scheme)**.** A *key-agreement scheme* KA is defined by the schema:

$$\begin{aligned} \mathsf{SecretKey} &: \mathsf{Type} \\ \mathsf{PublicKey} &: \mathsf{Type} \\ \mathsf{SharedSecret} &: \mathsf{Type} \\ \mathsf{derive} &: \mathsf{SecretKey} \to \mathsf{PublicKey} \\ \mathsf{agree} &: \mathsf{SecretKey} \times \mathsf{PublicKey} \to \mathsf{SharedSecret} \end{aligned}$$

with the following properties:

- **Agreement**: For all $\mathsf{sk}_1, \mathsf{sk}_2 : \mathsf{SecretKey}$, $\mathsf{agree}(\mathsf{sk}_1, \mathsf{derive}(\mathsf{sk}_2)) = \mathsf{agree}(\mathsf{sk}_2, \mathsf{derive}(\mathsf{sk}_1))$
- **TODO**: security properties

Definition 4.1.4 (Symmetric-Key Encryption Scheme)**.** A *symmetric-key encryption scheme* SYM is defined by the schema:

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

with the following properties:

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

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

- **TODO**: hiding, one-time encryption security?

Definition 4.1.5 (Key-Derivation Function)**.** A *key-derivation function* KDF defined over a symmetric-key encryption scheme SYM and a key-agreement scheme KA is a function of type:

$$\mathsf{KDF} : \mathsf{KA.SharedSecret} \to \mathsf{SYM.Key}$$

with the following properties:

- **TODO**: security properties

Definition 4.1.6 (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 $\mathsf{KA.SharedSecret}$ to $\mathsf{SYM.Key}$. We can define the following encryption and decryption algorithms:

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

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

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

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

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

Notation: We denote the above *encrypted message* type as $\mathsf{Message} := \mathsf{SYM.Ciphertext} \times \mathsf{KA.PublicKey}$, and the above two algorithms by

$$\begin{aligned}
&\mathsf{encrypt} : \mathsf{KA.SecretKey} \times \mathsf{KA.PublicKey} \times \mathsf{SYM.Plaintext} \to \mathsf{Message} \\
&\mathsf{decrypt} : \mathsf{KA.SecretKey} \times \mathsf{KA.PublicKey} \times \mathsf{SYM.Ciphertext} \to \mathsf{Option}(\mathsf{SYM.Plaintext})
\end{aligned}$$

TODO: security properties, combine with SYM and KA properties, like the fact that some of these keys are ephemeral, etc.

Definition 4.1.7 (Dynamic Cryptographic Accumulator)**.** A *dynamic cryptographic accumulator* DCA is defined by the schema:

$$\begin{aligned}
\mathsf{Item} &: \mathsf{Type} \\
\mathsf{Output} &: \mathsf{Type} \\
\mathsf{Witness} &: \mathsf{Type} \\
\mathsf{State} &: \mathsf{Type} \\
\mathsf{current} &: \mathsf{State} \to \mathsf{Output} \\
\mathsf{insert} &: \mathsf{Item} \times \mathsf{State} \to \mathsf{State} \\
\mathsf{contains} &: \mathsf{Item} \times \mathsf{State} \to \mathsf{Option}(\mathsf{Output} \times \mathsf{Witness}) \\
\mathsf{verify} &: \mathsf{Item} \times \mathsf{Output} \times \mathsf{Witness} \to \mathsf{Bool}
\end{aligned}$$

with the following properties:

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

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

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

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

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

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

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

- **TODO**: security properties

TODO: add finite capacity constraint, something like $\mathsf{insert} : \mathsf{Item} \times \mathsf{State} \to \mathsf{Option}(\mathsf{State})$ where it fails when capacity is reached

Definition 4.1.8 (Non-Interactive Zero-Knowledge Proving System)**.** A *non-interactive zero-knowledge proving* *system* NIZK is defined by the schema:

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

Notation: We use the following notation for a NIZK:

- We write the $\mathsf{Statement}$ and $\mathsf{ProvingKey}$ arguments of prove in the superscript and subscript respectively,

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

- We write the $\mathsf{VerifyingKey}$ argument of verify in the subscript,

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

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

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

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

- **Completeness**: For all $(x, w) : \mathsf{PublicInput} \times \mathsf{SecretInput}$, if $\mathsf{satisfying}^{P}_{\mathsf{pk}}(x, w) = \mathsf{True}$ with proof witness π, then $\mathsf{verify}_{\mathsf{vk}}(x, \pi) = \mathsf{True}$.
- **Knowledge Soundness**: For any polynomial-size adversary $\mathcal{A}$,

$$\mathcal{A} : \mathsf{ProvingKey} \times \mathsf{VerifyingKey} \to \mathfrak{D}(\mathsf{PublicInput} \times \mathsf{Proof})$$

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

$$\mathcal{E}_{\mathcal{A}} : \mathsf{ProvingKey} \times \mathsf{VerifyingKey} \to \mathfrak{D}(\mathsf{SecretInput})$$

  such that the following probability is negligible:

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

- **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:

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

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

4.2 Addresses and Key Components

Given a choice of HPKE we have the following definitions:

Definition 4.2.1 (Spending Key)**.** A $\mathsf{SpendingKey}$ is the following pair of keys:

$$\begin{aligned}
\mathsf{spend} &: \mathsf{HPKE.KA.SecretKey}\\
\mathsf{view} &: \mathsf{HPKE.KA.SecretKey}
\end{aligned}$$

The second secret key, view, is called the $\mathsf{ViewingKey}$.

Definition 4.2.2 (Receiving Key)**.** A $\mathsf{ReceivingKey}$ is the following pair of keys:

$$\begin{aligned}\mathsf{spend} &: \mathsf{HPKE.KA.PublicKey}\\ \mathsf{view} &: \mathsf{HPKE.KA.PublicKey}\end{aligned}$$

which is derived from a spending key $\mathsf{sk} : \mathsf{SpendingKey}$ with the following algorithm:

$$\begin{aligned}\mathsf{rk.spend} &:= \mathsf{KA.derive}(\mathsf{sk.spend})\\ \mathsf{rk.view} &:= \mathsf{KA.derive}(\mathsf{sk.view})\end{aligned}$$

A keypair $(\mathsf{sk}, \mathsf{rk}) : \mathsf{SpendingKey} \times \mathsf{ReceivingKey}$, represents the ability to spend and receive Assets as a unique *representative participant* on the Ledger. Any user of the MantaPay protocol can create many such keypairs, but each one represents a different participant and Assets must be transfered between them using the Transfer protocol as if they were independently owned by different users. A ReceivingKey can be used to receive any number of Assets and the SpendingKey can be used to spend any number of those Assets. See § 4.4 for the protocol used to spend a subset of Assets owned by a single user.

Important: To every spending key $\mathsf{sk} : \mathsf{SpendingKey}$ we have an assoicated viewing key $\mathsf{vk} : \mathsf{ViewingKey} := \mathsf{sk.view}$ which allows the owner to decrypt the encrypted messages associated to sk, but does not contain enough information to perform a spend with those Assets. This can be used for account auditing purposes, and for removing anonymity, but sharing this key should be done with caution.

In general, one may have a collection of viewing keys which can be used to separate the encrypted notes into different sets, by key. This way only certain transactions can be de-anonymized by certain parties.

4.3 Transfer Protocol

The Transfer protocol is the fundamental abstraction in MantaPay and facilitiates the valid transfer of Assets among participants while preserving their anonymity. The Transfer is made up of special cryptographic constructions called Senders and Receivers which represent the private input and the private output of a transaction. 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 ReceivingKeys from other participants for the outputs, assigning each key a subset of the input Assets, and then builds a Transfer object representing the transfer they want to build. From this Transfer object, they construct a TransferPost which they then send to the Ledger to be validated and stored, representing a completed state transition in the Ledger. 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 *must* 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:

- **Hybrid Public Key Encryption**: HPKE
- **UTXO Commitment Scheme**: $\mathsf{COM}^{\mathsf{UTXO}}$
- **Void Number Hash**: $\mathsf{HASH}^{\mathsf{VN}}$
- **Dynamic Cryptographic Accumulator**: DCA
- **Zero-Knowledge Proving System**: NIZK

with the following notational conventions:

$$\begin{aligned}\mathsf{KA} &:= \mathsf{HPKE.KA}\\ \mathsf{Trapdoor} &:= \mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Trapdoor}\\ \mathsf{UTXO} &:= \mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Output}\\ \mathsf{VoidNumber} &:= \mathsf{HASH}^{\mathsf{VN}}.\mathsf{Output}\\ \mathsf{EncryptedNote} &:= \mathsf{HPKE.Message}\\ \mathsf{UTXOSet} &:= \mathsf{DCA}\end{aligned}$$

and the following constraints:

$$\begin{aligned}
\mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Input} &= \mathsf{Asset} \\
\mathsf{HASH}^{\mathsf{VN}}.\mathsf{Input} &= \mathsf{UTXO} \times \mathsf{KA.SecretKey} \\
\mathsf{UTXOSet.Item} &= \mathsf{UTXO} \\
\mathsf{ValidTransfer} &: \mathsf{NIZK.Statement}
\end{aligned}$$

where ValidTransfer is defined below.

For the rest of this section, we assume the existence of a TransferConfiguration and use the primitives outlined above explicitly. 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:

$$\begin{aligned}
\mathsf{sk} &: \mathsf{SpendingKey} \\
\mathsf{epk} &: \mathsf{KA.PublicKey} \\
\mathsf{trapdoor} &: \mathsf{Trapdoor} \\
\mathsf{asset} &: \mathsf{Asset} \\
\mathsf{cm} &: \mathsf{UTXO} \\
\mathsf{cm}_z &: \mathsf{UTXOSet.Output} \\
\mathsf{cm}_w &: \mathsf{UTXOSet.Witness} \\
\mathsf{vn} &: \mathsf{VoidNumber}
\end{aligned}$$

A Sender, S, is constructed from a spending key sk : SpendingKey and an encrypted message note : EncryptedNote with the following algorithm:

$$\begin{aligned}
S.\mathsf{sk} &:= \mathsf{sk} \\
\mathsf{epk}, c &:= \mathsf{note} \\
\mathsf{Some}(\mathsf{asset}) &:= \mathsf{HPKE.decrypt}(S.\mathsf{sk.view}, \mathsf{epk}, c) \\
S.\mathsf{asset} &:= \mathsf{asset} \\
S.\mathsf{epk} &:= \mathsf{epk} \\
S.\mathsf{trapdoor} &:= \mathsf{KA.agree}(S.\mathsf{sk.spend}, S.\mathsf{epk}) \\
S.\mathsf{cm} &:= \mathsf{COM}^{\mathsf{UTXO}}(S.\mathsf{trapdoor}, S.\mathsf{asset}) \\
\mathsf{Some}(\mathsf{cm}_z, \mathsf{cm}_w) &:= \mathsf{UTXOSet.contains}(S.\mathsf{cm}, \mathsf{Ledger.utxos}()) \\
S.\mathsf{cm}_z &:= \mathsf{cm}_z \\
S.\mathsf{cm}_w &:= \mathsf{cm}_w \\
S.\mathsf{vn} &:= \mathsf{HASH}^{\mathsf{VN}}(S.\mathsf{cm}, S.\mathsf{sk.spend})
\end{aligned}$$

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

$$\begin{aligned}
\mathsf{cm}_z &: \mathsf{UTXOSet.Output} \\
\mathsf{vn} &: \mathsf{VoidNumber}
\end{aligned}$$

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:

$$\begin{aligned}
\mathsf{rk} &: \mathsf{ReceivingKey} \\
\mathsf{esk} &: \mathsf{KA.SecretKey} \\
\mathsf{trapdoor} &: \mathsf{Trapdoor} \\
\mathsf{asset} &: \mathsf{Asset} \\
\mathsf{cm} &: \mathsf{UTXO} \\
\mathsf{note} &: \mathsf{EncryptedNote}
\end{aligned}$$

A Receiver, R, is constructed from a receving key $\mathsf{rk} : \mathsf{ReceivingKey}$, an asset $\mathsf{asset} : \mathsf{Asset}$, and a random ephemeral secret key $\mathsf{esk} : \mathsf{HPKE.KA.SecretKey}$ with the following algorithm:

$$\begin{aligned}
R.\mathsf{rk} &:= \mathsf{rk} \\
R.\mathsf{esk} &:= \mathsf{esk} \\
R.\mathsf{trapdoor} &:= \mathsf{KA.agree}(R.\mathsf{esk}, R.\mathsf{rk.spend}) \\
R.\mathsf{asset} &:= \mathsf{asset} \\
R.\mathsf{cm} &:= \mathsf{COM}^{\mathsf{UTXO}}(R.\mathsf{trapdoor}, R.\mathsf{asset}) \\
R.\mathsf{note} &:= \mathsf{HPKE.encrypt}(R.\mathsf{rk.view}, R.\mathsf{esk}, R.\mathsf{asset})
\end{aligned}$$

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

$$\begin{aligned}
\mathsf{cm} &: \mathsf{UTXO} \\
\mathsf{note} &: \mathsf{EncryptedNote}
\end{aligned}$$

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:

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

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

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

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. It is not necessary to prove that the encryption of Receiver.note and the decryption of a note from the Ledger are valid. Deviation from the protocol in encryption or decryption stages does not reduce the security of the protocol for honest participants, it only makes certain assets inaccessible to honest receivers if they are not aware of the devation, since they cannot decrypt assets normally. This does not effect the *balanced transfer* or *ownership* invariants of the protocol for the existing assets of ledger participants.

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

1. All the AssetIds in T are equal:

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

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

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

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

$$\begin{aligned}
S.\mathsf{trapdoor} &= \mathsf{KA.agree}(S.\mathsf{sk.spend}, S.\mathsf{epk}) \\
S.\mathsf{cm} &= \mathsf{COM}^{\mathsf{UTXO}}(S.\mathsf{trapdoor}, S.\mathsf{asset}) \\
S.\mathsf{vn} &= \mathsf{HASH}^{\mathsf{VN}}(S.\mathsf{cm}, S.\mathsf{sk.spend}) \\
\mathsf{UTXOSet.verify}(S.\mathsf{cm}, S.\mathsf{cm}_z, S.\mathsf{cm}_w) &= \mathsf{True}
\end{aligned}$$

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

$$\begin{aligned}
R.\mathsf{note.epk} &= \mathsf{KA.derive}(R.\mathsf{esk}) \\
R.\mathsf{trapdoor} &= \mathsf{KA.agree}(R.\mathsf{esk}, R.\mathsf{rk.spend}) \\
R.\mathsf{cm} &= \mathsf{COM}^{\mathsf{UTXO}}(R.\mathsf{trapdoor}, R.\mathsf{asset})
\end{aligned}$$

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:

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

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

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

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 a participant has constructed a transfer post P : TransferPost they can send it to the Ledger for verification.

Definition 4.3.10 (Ledger-side Transfer Validity)**.** To check that P represents a valid Transfer, the ledger checks the following:

- **Public Withdraw**: All the public addresses corresponding to the Assets in P.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.sinks exist.
- **Current Accumulated State**: The UTXOSet.Output stored in each P.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.senders are unique, and no VoidNumber in P.senders has already been stored in the Ledger.VoidNumberSet.
- **New UTXOs**: All the UTXOs in P.receivers are unique, and no UTXO in P.receivers has already been stored on the ledger.
- **Verify Transfer**: Check that $\mathsf{NIZK.verify}_{\mathsf{vk}}(P.\mathsf{sources} \,||\, P.\mathsf{senders} \,||\, P.\mathsf{receivers} \,||\, P.\mathsf{sinks}, P.\pi) = \mathsf{True}$.

Definition 4.3.11 (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 Semantic 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 Semantic Transaction Algorithm

```
procedure BuildTransaction(sk, $\mathcal{B}$, total, rk)
    $B \leftarrow \mathsf{Sample}(\mathsf{total}, \mathcal{B})$        ▷ Samples pairs from $\mathcal{B}$ that total at least total
    if $\mathsf{len}(B) = 0$ then
        return [ ]        ▷ Insufficient Balance
    end if
    $P \leftarrow [\,]$        ▷ Allocate a new list for TransferPosts
    while $\mathsf{len}(B) > N$ do        ▷ While there are enough pairs to make another Transfer
        $A \leftarrow [\,]$
        for $b \in (B, N)$ do        ▷ Get the next $N$ pairs from $B$
            $S \leftarrow \mathsf{BuildSenders}_{\mathsf{sk}}(b)$
            $[acc, zs...] \leftarrow \mathsf{BuildAccumulatorAndZeroes}_{\mathsf{sk}}(S)$        ▷ Build a new accumulator and zeroes
            $P \leftarrow P + \mathsf{TransferPost}(\mathsf{Transfer}([], S, [acc, zs...], []))$
            $(A, Z) \leftarrow (A + (acc.\tilde{d}, acc.\mathsf{asset.value}), Z + zs)$        ▷ Save $acc$ for the next loop, $zs$ for the end
        end for
        $B \leftarrow A + \mathsf{remainder}(B, N)$
    end while
    $S \leftarrow \mathsf{PrepareZeroes}_{\mathsf{sk}}(N, B, Z, P)$        ▷ Use $Z$ and Mints to make $B$ go up to $N$ in size.
    $R \leftarrow \mathsf{BuildReceiver}_{\mathsf{sk}}(\mathsf{rk}, S)$
    $[c, zs...] \leftarrow \mathsf{BuildAccumulatorAndZeroes}_{\mathsf{sk}}(S)$
    return $P + \mathsf{TransferPost}(\mathsf{Transfer}([], S, [R, c, zs...], []))$
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 action that the user can take, performing a withdrawl from their shielded balance. To describe such a *semantic transaction*, we assume the existence of two transfer shapes[4]: 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, $\mathsf{sk} : \mathsf{SpendingKey}$, and asset id, $\mathsf{id} : \mathsf{AssetId}$, we are given a balance state, $\mathcal{B} : \mathsf{FinSet}\,(\mathsf{KA.PublicKey} \times \mathsf{AssetValue})$, a set of key-balance pairs for unspent assets, a total balance to withdraw, $\mathsf{total} : \mathsf{AssetValue}$, and a receiving key $\mathsf{rk} : \mathsf{ReceivingKey}$. We can then compute

$$\textsc{BuildTransaction}(\mathsf{sk}, \mathcal{B}, \mathsf{total}, \mathsf{rk})$$

to receive a List(TransferPost) to send to the ledger, representing the transfer of total to rk.

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 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.

[4]Other Transfer accumulation algorithms are possible with different starting shapes.

5 Concrete Protocol

TODO: other than cryptographic schemes, are there any implementation details we want to include here?

5.1 Concrete Cryptographic Schemes

Since we use a Zero-Knowledge Proving System, we want the cryptographic constructions below to be *ZKP-friendly*. The ZKP system we use is based on elliptic-curve cryptography so we will notate a fixed embedded elliptic curve group $\mathbb{G}$ with scalar field $\mathbb{F}$ for the following constructions. See Def 5.1.6 for more.

Definition 5.1.1 (Commitment Schemes and Hash Functions)**.** **TODO**: Poseidon hash for UTXO commitment (arity-4) or Pedersen Commitment variant and Poseidon hash for VN hash (arity-2)

Definition 5.1.2 (Key-Agreement Scheme)**.** For KA, we use a Diffie-Hellman Key Exchange over $(\mathbb{G}, \mathbb{F})$:

$$\mathsf{KA.derive}(x) : \mathbb{F} \to \mathbb{G} := x \cdot G$$
$$\mathsf{KA.agree}(x, y) : \mathbb{F} \times \mathbb{G} \to \mathbb{G} := x \cdot y$$

where G is a fixed public point.

Definition 5.1.3 (Symmetric-Key Encryption Scheme)**.** For SYM, we use

TODO: symmetric-key encryption scheme: AES-GCM with magic-number nonce and no associated data

Definition 5.1.4 (Key-Derivation Functions)**.** For KDF, we use ... **TODO**: Blake2s with magic-number salt

Definition 5.1.5 (Dynamic Cryptographic Accumulator)**.** For DCA, we use

TODO: dynamic cryptographic accumulator: Merkle Tree with Poseidon hashes (incremental tree for the ledger is an optimization since it only needs to know enough to compute the accumulated value)

Definition 5.1.6 (Non-Interactive Zero-Knowledge Proving System)**.** For NIZK, we use

TODO: non-interactive zero-knowledge proving system: Groth16 and/or PLONK

6 Acknowledgements

TODO: add acknowledgements

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.