MantaPay Protocol Specification

v0.5.0

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

April 23, 2022

Abstract

MantaPay is an implementation of a decentralized anonymous payment scheme based on the Mantapap protocol outlined in the original Manta whitepaper.

Contents

1	Introduction	2				
2	Notation	2				
3	Concepts 3.1 Assets 3.2 Addresses 3.3 Ledger 3.3.1 UTXOs and the UTXOSet 3.3.2 EncryptedNotes 3.3.3 VoidNumbers and the VoidNumberSet	2 2 3 4 4 4				
4	Abstract Protocol 4.1 Abstract Cryptographic Schemes 4.1.1 Commitment Scheme 4.1.2 Hash Function 4.1.3 Key-Agreement Scheme 4.1.4 Symmetric-Key Encryption Scheme 4.1.5 Key-Derivation Function 4.1.6 Hybrid Public Key Encryption Scheme 4.1.7 Dynamic Cryptographic Accumulator 4.1.8 Non-Interactive Zero-Knowledge Proving System 4.2 Addresses and Key Components 4.3 Transfer Protocol 4.4 Semantic Transactions	4 4 5 5 5 6 6 6 7 7 8 12				
5	Concrete Protocol 5.1 Poseidon Permutation 5.2 Elliptic Curve Cryptography 5.3 Concrete Cryptographic Schemes 5.3.1 Commitment Schemes 5.3.2 Key-Agreement Scheme 5.3.3 Symmetric-Key Encryption Scheme 5.3.4 Key-Derivation Functions 5.3.5 Dynamic Cryptographic Accumulator 5.3.6 Non-Interactive Zero-Knowledge Proving System	13 13 13 13 13 13 13 13 13				
6	Acknowledgements 1					
7	References	14				

 $^{^{*}}$ 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	X
Monero	RingCT/NIZK	PoW	1	Х
Tornado Cash (Nova)	NIZK	Х	2	✓
MantaPay 0.5.0	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 types¹.
- If x:T then x is a value and T is a type, denoted T: Type, and we say that x has type T.
- Bool is the type of booleans with values True and False.
- For any types A: Type and B: Type we denote the type of functions from A to B as $A \to B$: Type.
- For any types A: Type and B: Type we denote the *product type* over A and B as $A \times B$: Type with constructor $(-,-):A \to (B \to A \times B)$. Depending on context, we may omit the constructor and inline the pair into another constructor/destructor. For example, if $f:A \times B \to C$ we can denote f((a,b)) as f(a,b) to reduce the number of parentheses.
- For any type T: Type, we define $\mathsf{Option}(T)$: Type as the inductive type with constructors:

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

- We denote the *type of finite sets* over a type T: Type as $\mathsf{FinSet}(T)$: 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: Type as $\mathsf{List}(T)$: 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 $[\dots]$ for an arbitrary set of elements.
- We denote the type of distributions over a type T: Type as $\mathfrak{D}(T)$: 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, $y \in \{x \mid x \sim \mathfrak{D}(T)\} \leftrightarrow y \in \mathfrak{D}(T)$.
- 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: 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.

¹By 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.

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 AssetId². 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.

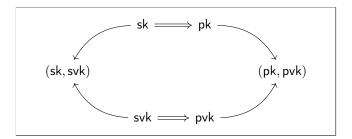


Figure 1: Key pairs and Addresses.

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 (pk, 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 ledger³ that keeps track of the movement of Assets from one participant to another. Unfortu-

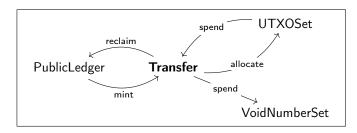


Figure 2: Lifecycle of an Asset.

nately, 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

²It 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*.

³A 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.

as follows: the public ledger as PublicLedger, the shielded asset pool as ShieldedAssetPool, and the combined ledger we denote Ledger.

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, 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:

Randomness: Type

Input : Type

Output: Type

 $RandomnessDistribution : \mathfrak{D}(Randomness)$

 $\mathsf{commit} : \mathsf{Randomness} \times \mathsf{Input} \to \mathsf{Output}$

with the following properties:

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

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

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

Input : Type Output : Type

 $\mathsf{hash}:\mathsf{Input}\to\mathsf{Output}$

with the following properties:

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

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

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

SecretKey : Type PublicKey : Type SharedSecret : Type

 $\mathsf{derive} : \mathsf{SecretKey} \to \mathsf{PublicKey}$

 $agree : SecretKey \times PublicKey \rightarrow SharedSecret$

with the following properties:

- Agreement: For all sk_1, sk_2 : SecretKey, $agree(sk_1, derive(sk_2)) = agree(sk_2, derive(sk_1))$
- Passive Security: Even if an adversary eavedrops on the network communication, she cannot forge the agreed secret unless she knows how to find a preimage for 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 4.1.4 (Symmetric-Key Encryption Scheme). A one-time symmetric-key encryption scheme SYM is defined by the schema:

Key : Type Plaintext : Type Ciphertext : Type

 $\mathsf{encrypt} : \mathsf{Key} \times \mathsf{Plaintext} \to \mathsf{Ciphertext}$

 $decrypt : Key \times Ciphertext \rightarrow Option(Plaintext)$

with the following properties:

• Soundness: For all keys k: Key and plaintexts p: Plaintext, we have that

$$decrypt(k, encrypt(k, p)) = Some(p)$$

• Security Requirement: The symmetric-key encryption scheme must be one-time (INT-CTXT ∧ 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 4.1.5 (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}.\mathsf{SharedSecret} \to \mathsf{SYM}.\mathsf{Key}
```

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 KA.SharedSecret to SYM.Key. We can define the following encryption and decryption algorithms:

• Encryption: Given an ephemeral secret key esk : KA.SecretKey, a public key pk : KA.PublicKey, and plaintext p: SYM.Plaintext, we produce the pair

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

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

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

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

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

```
\label{eq:continuous_encrypt} \begin{split} &\mathsf{encrypt}: \mathsf{KA}.\mathsf{SecretKey} \times \mathsf{KA}.\mathsf{PublicKey} \times \mathsf{SYM}.\mathsf{Plaintext} \to \mathsf{Message} \\ &\mathsf{decrypt}: \mathsf{KA}.\mathsf{SecretKey} \times \mathsf{KA}.\mathsf{PublicKey} \times \mathsf{SYM}.\mathsf{Ciphertext} \to \mathsf{Option}(\mathsf{SYM}.\mathsf{Plaintext}) \end{split}
```

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

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

Item: Type
Output: Type
Witness: Type
State: Type

 $\mbox{current}: \mbox{State} \rightarrow \mbox{Output} \\ \mbox{insert}: \mbox{Item} \times \mbox{State} \rightarrow \mbox{State} \\$

contains : Item \times State \rightarrow Option(Output \times Witness)

 $verify: Item \times Output \times Witness \rightarrow Bool$

with the following properties:

• Unique Accumulated Values: For any initial state s: State and any list of items I: List(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: State and any list of items I: List(Item) we can generate the sequences of states:

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

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

$$Some(z, w) := contains(t, s), verify(t, z, w) = True$$

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

Statement : Type ProvingKey : Type VerifyingKey : Type PublicInput : Type SecretInput : Type Proof : Type

keys : Statement $\rightarrow \mathfrak{D}(\mathsf{ProvingKey} \times \mathsf{VerifyingKey})$

 $\mathsf{prove}: \mathsf{Statement} \times \mathsf{ProvingKey} \times \mathsf{PublicInput} \times \mathsf{SecretInput} \to \mathfrak{D}(\mathsf{Option}(\mathsf{Proof}))$

 $\mathsf{verify}: \mathsf{VerifyingKey} \times \mathsf{PublicInput} \times \mathsf{Proof} \to \mathsf{Bool}$

Notation: We use the following notation for a NIZK:

• We write the Statement and 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 VerifyingKey argument of verify in the subscript,

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

• We say that (x, w): PublicInput \times SecretInput has the property of being a 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: Statement and keys (pk, vk) $\sim \text{keys}(P)$:

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

$$\Pr \begin{bmatrix} \mathsf{satisfying}_{\mathsf{pk}}^P(x,w) = \mathsf{False} \\ \mathsf{verify}_{\mathsf{vk}}(x,w) = \mathsf{True} \end{bmatrix} \begin{pmatrix} (\mathsf{pk},\mathsf{vk}) \sim \mathsf{keys}(P) \\ (x,\pi) \sim \mathcal{A}(\mathsf{pk},\mathsf{vk}) \\ w \sim \mathcal{E}_{\mathcal{A}}(\mathsf{pk},\mathsf{vk}) \end{bmatrix}$$

• Statistical Zero-Knowledge: There exists a stateful simulator S, such that for all stateful distinguishers D, the difference between the following two probabilities is negligible:

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

• Succinctness: For all (x, w): PublicInput \times 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 SpendingKey is the following pair of keys:

spend : HPKE.KA.SecretKey
view : HPKE.KA.SecretKey

The second secret key, view, is called the ViewingKey.

Definition 4.2.2 (Receiving Key). A Receiving Key is the following pair of keys:

spend : HPKE.KA.PublicKey view : HPKE.KA.PublicKey

which is derived from a spending key sk: SpendingKey with the following algorithm:

rk.spend := KA.derive(sk.spend)rk.view := KA.derive(sk.view)

A keypair (sk, rk): SpendingKey \times 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 \S 4.4 for the protocol used to spend a subset of Assets owned by a single user.

Important: To every spending key sk: SpendingKey we have an assoicated viewing key vk: ViewingKey := 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:

 \bullet Hybrid Public Key Encryption: HPKE

• UTXO Commitment Scheme: COM^{UTXO}

• Void Number Commitment Scheme: COM^{VN}

 \bullet Dynamic Cryptographic Accumulator: DCA

• Zero-Knowledge Proving System: NIZK

with the following notational conventions:

 $\mathsf{KA} := \mathsf{HPKE}.\mathsf{KA}$ $\mathsf{UTXO} := \mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Output}$ $\mathsf{VoidNumber} := \mathsf{COM}^{\mathsf{VN}}.\mathsf{Output}$

EncryptedNote := HPKE.Message

 $\mathsf{UTXOSet} := \mathsf{DCA}$

and the following constraints:

$$\begin{split} \mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Randomness} &= \mathsf{KA}.\mathsf{SecretKey} \\ &\quad \mathsf{COM}^{\mathsf{UTXO}}.\mathsf{Input} &= \mathsf{KA}.\mathsf{PublicKey} \times \mathsf{Asset} \\ &\quad \mathsf{COM}^{\mathsf{VN}}.\mathsf{Randomness} &= \mathsf{KA}.\mathsf{SecretKey} \\ &\quad \mathsf{COM}^{\mathsf{VN}}.\mathsf{Input} &= \mathsf{UTXO} \\ &\quad \mathsf{UTXOSet}.\mathsf{Item} &= \mathsf{UTXO} \end{split}$$

ValidTransfer: NIZK.Statement

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:

sk : SpendingKey

rk.spend: KA.PublicKey

 $\mathsf{esk}: \mathsf{KA}.\mathsf{SecretKey}$

 $\mathsf{asset} : \mathsf{Asset}$

cm : UTXO

 $\mathsf{cm}_z:\mathsf{UTXOSet}.\mathsf{Output}$

 $\mathsf{cm}_w : \mathsf{UTXOSet.Witness}$

vn: VoidNumber

A Sender, S, is constructed from a spending key sk : SpendingKey and an encrypted message (epk, C_{note}) : EncryptedNote with the following algorithm:

 $S.\mathsf{sk} := \mathsf{sk}$

S.rk.spend := KA.derive(S.sk.spend)

 $\mathsf{Some}\left(S.\mathsf{esk}, S.\mathsf{asset}\right) := \mathsf{HPKE}.\mathsf{decrypt}(S.\mathsf{sk}.\mathsf{view}, \mathsf{epk}, C_{\mathsf{note}})$

 $S.\mathsf{cm} := \mathsf{COM}^{\mathsf{UTXO}}_{S.\mathsf{esk}}(S.\mathsf{rk.spend}, S.\mathsf{asset})$

 $Some(S.cm_z, S.cm_w) := UTXOSet.contains(S.cm, Ledger.utxos())$

 $S.vn := COM_{S.sk.spend}^{VN}(S.cm)$

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

 cm_z : UTXOSet.Output vn: VoidNumber

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:

rk: ReceivingKey

esk : KA.SecretKey

asset : Asset cm : UTXO

note: Encrypted Note

A Receiver, R, is constructed from a receiving key rk: Receiving Key, an asset asset: Asset, and a random ephemeral secret key esk: HPKE.KA.Secret Key with the following algorithm:

 $R.\mathsf{rk} := \mathsf{rk}$ $R.\mathsf{esk} := \mathsf{esk}$ $R.\mathsf{asset} := \mathsf{asset}$ $R.\mathsf{cm} := \mathsf{COM}^{\mathsf{UTXO}}_{R.\mathsf{esk}}(R.\mathsf{rk}.\mathsf{spend}, R.\mathsf{asset})$ $R.\mathsf{note} := \mathsf{HPKE}.\mathsf{encrypt}\left(R.\mathsf{rk}.\mathsf{view}, R.\mathsf{esk}, (R.\mathsf{esk}, R.\mathsf{asset})\right)$

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

cm : UTXO note : EncryptedNote

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:

sources : List(Asset)
senders : List(Sender)
receivers : List(Receiver)
sinks : List(Asset)

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

$$(|T.sources|, |T.senders|, |T.receivers|, |T.sinks|)$$

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. 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: Transfer is considered valid if and only if

1. All the AssetIds in T are equal:

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

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

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

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

$$S.\mathsf{rk.spend} = \mathsf{KA.derive}(S.\mathsf{sk.spend})$$

$$S.\mathsf{cm} = \mathsf{COM}^{\mathsf{UTXO}}_{S.\mathsf{esk}}(S.\mathsf{rk.spend}, S.\mathsf{asset})$$

$$S.\mathsf{vn} = \mathsf{COM}^{\mathsf{VN}}_{S.\mathsf{sk.spend}}(S.\mathsf{cm})$$

$$\mathsf{UTXOSet.verify}(S.\mathsf{cm}, S.\mathsf{cm}_z, S.\mathsf{cm}_w) = \mathsf{True}$$

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

$$R.\mathsf{cm} = \mathsf{COM}^{\mathsf{UTXO}}_{R.\mathsf{esk}}(R.\mathsf{rk.spend}, R.\mathsf{asset})$$

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{tabular}{ll} sources : List(Source) \\ senders : List(SenderPost) \\ receivers : List(ReceiverPost) \\ sinks : List(Sink) \\ \hline $\pi : NIZK.Proof$ \\ \end{tabular}$

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

```
x \coloneqq \mathsf{Transfer.public}(T) w \coloneqq \mathsf{Transfer.secret}(T) \mathsf{Some}(\pi) \sim \mathsf{NIZK.prove}_{\mathsf{pk}}^{\mathsf{ValidTransfer}}(x, w) P.\mathsf{sources} \coloneqq x.\mathsf{sources} P.\mathsf{senders} \coloneqq x.\mathsf{senders} P.\mathsf{receivers} \coloneqq x.\mathsf{receivers} P.\mathsf{sinks} \coloneqq x.\mathsf{sinks} P.\pi \coloneqq \pi
```

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 NIZK.verify_{vk}(P.sources || P.senders || P.receivers || P.sinks, $P.\pi$) = 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, B, total, rk)
    B \leftarrow \mathsf{Sample}(\mathsf{total}, \mathcal{B})
                                                                                  \triangleright Samples pairs from \mathcal{B} that total at least total
    if len(B) = 0 then
                                                                                                                     ▷ Insufficient Balance
         return []
    end if
    P \leftarrow []
                                                                                              ▷ Allocate a new list for TransferPosts
    while len(B) > N do
                                                                     ▶ While there are enough pairs to make another Transfer
         A \leftarrow []
         for b \in (B, N) do
                                                                                                       \triangleright Get the next N pairs from B
             S \leftarrow \mathsf{BuildSenders}_{\mathsf{sk}}(b)
             [acc, zs...] \leftarrow \mathsf{BuildAccumulatorAndZeroes_{sk}}(S)
                                                                                              ▶ Build a new accumulator and zeroes
              P \leftarrow P + \mathsf{TransferPost}(\mathsf{Transfer}([], S, [acc, zs...], []))
             (A, Z) \leftarrow (A + (acc.d, acc.asset.value), Z + zs)
                                                                                       \triangleright 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)
                                                                              \triangleright Use Z and Mints to make B go up to N in size.
    R \leftarrow \mathsf{BuildReceiver_{sk}(rk}, S)
    [c, zs...] \leftarrow \mathsf{BuildAccumulatorAndZeroes_{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 shapes⁴: 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, sk: SpendingKey, and asset id, id: AssetId, we are given a balance state, \mathcal{B} : FinSet (KA.PublicKey \times AssetValue), a set of key-balance pairs for unspent assets, a total balance to withdraw, total: AssetValue, and a receiving key rk: ReceivingKey. We can then compute

```
BUILDTRANSACTION(sk, \mathcal{B}, total, 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.

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

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 protocol and reducing the overall cost of the Zero-Knowledge circuits. **Poseidon** (without sponges) is a family of hash functions with the following signature:

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

over some sufficiencly large finite field \mathbb{F} . 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: $\mathbb{S} \to \mathbb{F}$ to lift scalars to the proof system field.

5.3 Concrete Cryptographic Schemes

Definition 5.3.1 (Commitment Schemes). The protocol features two different commitment schemes: COM^{UTXO} the UTXO Commitment Scheme and COM^{VN} the Void Number Commitment Scheme. Both commitment schemes use **Poseidon** as the cryptographic primitive. The UTXO uses an arity-4 **Poseidon** with the following mapping:

$$COM^{UTXO}(esk, pk, asset) := Poseidon_4(lift(esk), x(pk), asset.id, asset.value)$$

where x is the coordinate-extraction function from \mathbb{G} to \mathbb{F} . For the Void Number Commitment Scheme we use an arity-2 **Poseidon** with the following mapping:

$$\mathsf{COM}^{\mathsf{VN}}(\mathsf{sk},\mathsf{cm}) := \mathbf{Poseidon}_2(\mathsf{lift}(\mathsf{sk}),\mathsf{cm})$$

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

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

where G is a fixed public point.

Definition 5.3.3 (Symmetric-Key Encryption Scheme). For SYM, we use symmetric-key encryption scheme: AES-GCM [8] with magic-number nonce and no associated data. Note that it is safe to reuse the nonce here because we assume that the encryption key is only used once. See Def 4.1.4 for more.

Definition 5.3.4 (Key-Derivation Functions). For KDF, we use Blake2s [7] with magic-number salt.

Definition 5.3.5 (Dynamic Cryptographic Accumulator). For DCA, we use a Merkle Tree with **Poseidon**₂ 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 5.3.6 (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].

6 Acknowledgements

We would like to thank Luke Pearson and Toghrul Maharramov 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 plookup. *IACR Cryptol. ePrint Arch.*, page 86, 2022.
- [7] Markku-Juhani O. Saarinen and Jean-Philippe Aumasson. The BLAKE2 Cryptographic Hash and Message Authentication Code (MAC). IETF-RFC 7693, Internet Engineering Task Force, 2015.
- [8] Joseph Salowey, Abhijit Choudhury, and David McGrew. AES Galois Counter Mode (GCM) Cipher Suites for TLS. IETF-RFC 5288, Internet Engineering Task Force, 2008.