MProve-Nova: A Privacy-Preserving Proof of Reserves Protocol for Monero

Varun Thakore, Saravanan Vijayakumaran
Department of Electrical Engineering, Indian Institute of Technology Bombay
Mumbai, Maharashtra, India
varunt@ee.iitb.ac.in,sarva@ee.iitb.ac.in

ABSTRACT

We present MProve-Nova, a proof of reserves (PoR) protocol for Monero that leverages the Nova recursive SNARK to achieve two firsts (without requiring any trusted setup). It is the first privacy-preserving Monero PoR protocol that does not reveal any information about the exchange-owned outputs or their key images. It is the first Monero PoR protocol where the proof size and proof verification time are *constant*, i.e. they are *independent* of the number of outputs on the Monero blockchain and the number of outputs owned by the exchange.

In our implementation of MProve-Nova, we observed proof sizes of 27 KB and verification times of 4.5 seconds. Proving times increase linearly with the number of outputs owned by the exchange (≈ 0.7 hour per 1,000 outputs) but remain independent of the number of outputs on the Monero blockchain. We also describe the design and implementation of a Nova-based non-collusion protocol that takes the MProve-Nova proofs from a pair of exchanges as input and proves that the sets of outputs they used to generate their respective proofs of reserves are non-overlapping.

KEYWORDS

Cryptocurrency, Monero, Proof of Reserves, Nova.

1 INTRODUCTION

Cryptocurrency exchanges provide a user-friendly platform for buying, selling, and trading of cryptocurrencies. While customers can transfer their coins from exchanges to non-custodial wallets, many of them prefer to keep their coins on exchanges to avoid the hassles and risks of managing private keys. This leaves customer funds at the risk of being stolen from the exchange due to security breaches or internal fraud.

Some examples include the bankruptcy of Mt. Gox [50] and the collapse of FTX [49]. The total funds lost due to exchange hacks alone is estimated to be at least $2.4 billion as of April 2023 [12]. While losses due to security breaches can be avoided by hardening the protocols involving the exchange's private keys, internal fraud by exchange operators cannot be completely prevented. But such fraud can be detected early (and hence deterred) if exchanges are required to regularly publish *proofs of solvency*.

A *proof of reserves (PoR)* is one half of a proof of solvency protocol, with a *proof of liabilities (PoL)* being the other half. A PoR protocol enables an exchange to prove that it owns a certain amount of a cryptocurrency, i.e. it holds a certain amount of assets. For this reason, a PoR is sometimes also called a *proof of assets*. A PoL protocol enables an exchange to prove that the total amount of assets it is storing on behalf of all its customers (its liabilities) equals a certain amount. If an exchange's assets exceeds its liabilities, it is considered solvent.

In this paper, we focus solely on PoR protocols for Monero [34], a privacy-focused cryptocurrency based on the CryptoNote protocol [47]. Specifically, we are interested in *publicly-verifiable privacy-preserving* PoR protocols. By publicly-verifiable, we mean that the PoR can be verified by any party, not just a trusted auditor. By privacy-preserving, we mean that the protocols do not reveal the specific coin addresses (*outputs*) owned by the exchange.

Without this privacy requirement, it is trivial to construct a PoR protocol for Monero. The exchange can generate signatures proving ownership of a set of outputs and non-membership proofs proving that the outputs have not been spent. The latter would involve proving that the outputs' *key images*[1] have not appeared in any past transaction.

Privacy is especially important in Monero PoR protocols because Monero transactions contain ring signatures, with the output being spent hidden among decoy outputs sampled from the Monero blockchain. If some outputs are identified as *unspent* outputs belonging to an exchange, they can be marked as decoys in all the previous transaction rings they appear in. This reduces the effective ring size of such transactions. The implication is that a non-private Monero PoR protocol can negatively impact the privacy of other Monero users, in addition to impacting the privacy of the exchange generating the proof.

When the identities of an exchange's outputs are hidden by a PoR protocol, it opens up the possibility of *collusion* between exchanges. Collusion refers to the situation when the same output is used by two exchanges to generate their respective proofs of reserves. This is a form of double spending, where one exchange could bribe another to contribute to the former's PoR. It is desirable to have privacy-preserving PoR protocols that are also *collusion-resistant*.

Finally, a proof of solvency is useless if no one verifies it. So it is desirable to have PoR/PoL protocols with short proofs that can be verified quickly on personal computers. This will lower the bar for proof verifiers, make it more likely that the proofs of reserves are verified by customers, and hence reduce the likelihood of exchanges in engaging in activities that could render them insolvent.

1.0.1 Paper organization. In the next section, we describe the challenges involved in designing Monero PoR protocols. Our contributions are listed in Section 3. Section 4 describes related work. In Section 5, we briefly cover aspects of Nova background required to describe MProve-Nova. Section 6 describes the design of the

[1] Key images are one-way functions of outputs that are generated as part of the Monero ring signatures to prevent double spending.

MProve-Nova protocol and the proof of non-collusion. We provide implementation details and performance results in Section 8 followed by conclusions in Section 9.

2 CHALLENGES IN DESIGNING A MONERO POR PROTOCOL

In this section, we describe the challenges involved in designing Monero PoR protocols. We also describe how previous PoR protocols for Monero address these challenges and where they fall short. This description will help clarify our contributions listed in Section 3. Readers unfamiliar with Monero can consult Appendix A to understand the terminology used in the rest of the paper.

Many cryptocurrency protocols (including Bitcoin) have the notion of an *unspent transaction output (UTXO)*. As the name suggests, this corresponds to an output having coins that have not been spent by their owner. For such cryptocurrencies, a PoR protocol can restrict its attention to the UTXO set and ignore all spent transaction outputs.

Since Monero hides the identity of the spending key in a transaction, one *cannot* partition the output set into spent and unspent outputs. While some transaction graph analysis techniques have been able to categorize a large percentage of non-RingCT outputs as spent [27, 36], it is not possible to know if any of the RingCT outputs (except for 5 of them) have been spent [48, 51].

Any Monero PoR protocol must prove two compound statements:

1. The prover knows the private keys corresponding to some outputs on the Monero blockchain. The sum of the coins in the output commitments will add up to the reserves owned by the prover.
2. The outputs contributing to the prover's reserves have not been already spent in a past transaction.

As discussed in the previous section, one cannot simply reveal the prover-owned outputs as this would violate the privacy of both the prover and other Monero users. One solution is to give a privacy-preserving proof of membership of prover-owned outputs in the set of all outputs, together with a proof of knowledge of private keys for each such output.

But this is not enough. One needs to also give a privacy-preserving proof that the prover-owned outputs are *unspent*. Furthermore, one needs to ensure that a malicious prover does not *double count*, i.e. it does not artificially inflate its reserves by counting the coins in an output more than once.

Previous Monero PoR protocols, namely MProve [20] and MProve+ [19], partially achieved the above goals as follows:

- They both proved that the prover-owned outputs belonged to an *anonymity set*, a subset of the set of all outputs that appeared on the Monero blockchain. For performance reasons, they chose the anonymity set to be much smaller than the set of all outputs.
- They both revealed one key image per prover-owned output that contributed to the reserves. This had several implications:
  - The verifier could verify that the prover was using only unspent outputs by checking that the revealed key images had not appeared in any transaction on Monero blockchain.
  - A malicious prover could not double count, as using an output more than once in the same PoR would result in a repeated key image.
  - The verifier could verify that two different provers were not colluding by checking that the sets of key images revealed by them were non-overlapping. So the proof of non-collusion was implicitly available in the original protocol itself.

But these approaches had some drawbacks.

- The proof size and proof verification time in both MProve and MProve+ increased linearly with the size of the anonymity set. Additionally, in MProve+ the memory required for generating and verifying proofs also increased linearly with the size of the anonymity set.
- More seriously, revealing the key images negatively impacted the privacy of the prover and, in some cases, the privacy of regular Monero users.
  - In MProve, if the prover spent from an address that was used to previously generate a MProve proof, then the address was immediately identified as the spending address in the transaction ring. This meant that all previous transaction rings where this address appeared as a decoy had their effective ring size reduced by one.
  - MProve+ improved upon MProve by breaking the direct link between the key image and a prover-owned address. If the prover spent from an address that was used to previously generate a MProve+ proof, then that address could be identified as the spending address only if the transaction ring and the anonymity set had exactly one address in common. But the transaction would always be identified as being initiated by the prover.

In this paper, we propose a Monero PoR protocol which does not suffer these drawbacks, while preventing malicious provers from using spent outputs or double counting unspent outputs.

3 OUR CONTRIBUTIONS

Our contributions are as follows:

1. We describe the design and implementation of MProve-Nova, a publicly-verifiable privacy-preserving PoR protocol for Monero. Our design is based on Nova [26], a recursive SNARK that does not require a trusted setup. MProve-Nova has the following advantages:
   (a) It is privacy-preserving in the random oracle model in the sense that it does not reveal any information about the exchange-owned outputs or their key images.
   (b) It is the first Monero PoR protocol with *constant* proof size and proof verification time i.e. these two metrics are *independent* of the number of outputs on the Monero blockchain and the number of outputs owned by the exchange. It achieves this by leveraging the Nova folding scheme to fold all the individual statements proving ownership and ”unspent-ness” of each of the exchange-owned outputs into a single statement.

(2) We provide a formal analysis of the security properties of MProve-Nova. We show that it is
   (a) *inflation resistant*, i.e. a computationally bounded exchange cannot generate an MProve-Nova proof that proves that it owns a number of coins that is greater than the actual number of coins it owns, and
   (b) *privacy-preserving in the random oracle model*
(3) We modified the reference implementation of Nova [32] to add zero-knowledge. The existing implementation is not zero-knowledge as indicated by an open [33] issue on their repository. We referred to another implementation of Nova [3] which added zero-knowledge in an older version of the reference implementation. We added zero-knowledge by implementing hiding commitments, hiding the number of steps executed and using zero-knowledge sumcheck and hiding inner product argument. We used our modified implementation of Nova [44] to implement MProve-Nova in Rust [37]. To leverage Nova, the statements requiring proof need to be expressed using rank-1 constraint system (R1CS) constraints. While it was possible to create an R1CS gadget for MProve-Nova using the circom circuit compiler [22] and the Nova Scotia adapter [9], we chose to implement the gadget directly using the bellpepper library [29].
   To implement the MProve-Nova gadget, we implemented the following component gadgets in bellpepper.
   (a) *Regular and indexed Merkle trees* [4, 45]: While implementations of regular Merkle trees already existed in bellman [53] (which was forked to create bellpepper), our implementation of indexed Merkle trees is new.
   (b) bellpepper-emulated: A gadget for non-native finite field arithmetic inspired by the emulated [17] package (written in Go) from the gnark zk-SNARK library [10]. When the field order is a pseudo-Mersenne prime, we added limb folding and efficient field membership checks inspired by techniques from circom-ecdsa [1].
   (c) bellpepper-ed25519: A gadget for Ed25519 elliptic curve operations using bellperson-emulated. This was needed as Monero uses Ed25519 curve points for addresses (outputs) and amount commitments.
(4) We describe the design, implementation [37], and security properties of a privacy-preserving protocol for generating a *proof of non-collusion* between a pair of exchanges that use MProve-Nova to generate their respective proofs of reserves. This protocol is also based on Nova, resulting in constant proof sizes and proof verification times.

4 RELATED WORK

Provisions [18], proposed by Dagher *et al.*, is one of the first privacy-preserving proof of solvency protocol for Bitcoin exchanges. It consists of three sub-protocols: proof of reserves, proof of liabilities and proof of solvency. In proof of reserves, the exchange generates a Pedersen commitment [40] C_{res} to the total assets corresponding to a subset of owned addresses $\mathcal{P}_{\text{own}}$ from a larger anonymity set $\mathcal{P}$. The exchange submits a proof that it included only those amounts in C_{res} for which it knows the private keys corresponding to the addresses in $\mathcal{P}$. In proof of liabilities, the exchange generates a Pedersen commitment to each bit of the balance amount owned by the customer. These commitments are combined to calculate the total liabilities $C_{\text{liabilities}}$ of the exchange. In proof of solvency, the exchange computes $C_{\text{diff}} = C_{\text{res}} - C_{\text{liabilities}}$ and proves that C_{diff} commits to a non-negative amount. There is also a fourth protocol to prove non-collusion but it reveals the number of addresses owned by the exchange and is presented as an optional protocol. Provisions is specific to Bitcoin and cannot be applied to privacy-preserving cryptocurrencies such as Monero.

In 2019, Blockstream [41, 42] released a tool to generate proof of reserves which involves generating an invalid transaction using all the UTXOs of an exchange and an invalid input so that exchange's funds are not spent. This technique does not preserve address privacy since all the UTXOs owned by the exchange are revealed.

Stoffu Noether implemented a technique for generating proof of reserves for Monero which was added to the official Monero client in 2018 [43]. It takes a target amount as input and finds the smallest set of addresses owned by the prover whose total amount exceeds the target amount. Then the set of addresses and their corresponding key images are revealed as part of the proof of reserves. This technique reveals the addresses owned by the prover, their corresponding amounts, and their key images.

Dutta *et al.* [20] proposed MProve, a proof of reserves protocol for Monero exchanges. MProve+ [19] was later proposed which enhanced the privacy of MProve by using techniques from Bulletproofs [11] and Omniring [30]. The details of MProve and MProve+ with drawbacks have been covered in Section 2.

Some notable work on proof of reserves include, gOTzilla [7] proposed by Baldimtsi *et al.*, which is an interactive zero-knowledge protocol for proving disjunctive statements. Additionally, Chatzigiannis and Chalkias [16] proposed proof of assets for account-based blockchains such as Diem, formerly known as Libra [2].

There has been significant work in proof of liabilities (PoL), most notable ones being DAPOL [14] and DAPOL+ [23]. MProve-Nova can be used with DAPOL+ to provide a proof of solvency. Chalkias *et al.* [13] highlighted vulnerabilities in the implementation of PoL used in production. Chatzigiannis *et al.* [15] evaluated and systematized several distributed payment systems which offer auditability. Their work provides a comparison between different proof of assets and proof of liabilities schemes on the basis of their efficiency and privacy properties.

5 NOVA

In this section, we cover aspects of the Nova recursive SNARK [26] required to describe the MProve-Nova protocol. Nova introduced a non-interactive folding scheme for committed relaxed rank-1 constraint systems (R1CS) [26, Construction 2]. It consists of two main components: an *incrementally verifiable computation (IVC)* [46] scheme and a zkSNARK to prove knowledge of valid IVC proofs.

5.0.1 IVC Scheme. An IVC scheme allows a prover to prove that for some function F and public values z_0 and z_n, it knows private inputs $w_0, w_1, \ldots, w_{n-1}$ such that

$$z_n = F\left(F\left(\ldots F\left(F\left(F\left(z_0, w_0\right), w_1\right), w_2\right), \ldots\right), w_{n-1}\right).$$

As shown in Fig. 1, this is achieved by proving the execution of a series of incremental computations of the form $z_{i+1} = F(z_i, w_i)$, for each $i \in \{0, 1, \ldots, n-1\}$, where z_i and z_{i+1} are the public input and output in the ith step, respectively.

The Nova IVC scheme uses a non-interactive folding scheme for committed relaxed R1CS. The step function F needs to be expressed using R1CS constraints. At each step i, the variables z_i, z_{i+1}, and w_i define the committed relaxed R1CS instance. This instance is folded into a running committed relaxed R1CS which represents the correct execution of steps $0, 1, \ldots, i-1$.

The IVC prover gives a proof Π_{i+1} at each step i, which attests that $z_{i+1} = F(z_i, w_i)$ was computed correctly and the folding of the two committed relaxed R1CS instances is valid. The IVC proof Π_{i+1} attests to the correct execution of steps $0, 1, \ldots, i$.

The Nova IVC scheme satisfies completeness and knowledge soundness. For more details on the Nova IVC scheme, we refer the reader to Section 5 of the Nova paper [26].

5.0.2 zkSNARK of IVC Proof. After n steps, the IVC prover produces a proof Π_n that attests to the correct execution of steps $0, 1, \ldots, n-1$. The IVC prover can send this proof to the verifier, but this does not satisfy zero-knowledge since the proof Π_n does not hide the prover's private inputs.

Instead, Nova uses a zero-knowledge SNARK (zkSNARK) to prove knowledge of a valid IVC proof Π_n.

(1) The prover $\mathcal{P}_{zk}$ and verifier $\mathcal{V}_{zk}$ of zkSNARK are given the instance (n, z_0, z_n).
(2) The prover $\mathcal{P}_{zk}$ additionally takes the proving key pk and IVC proof Π_n to produce the proof $\pi \longleftarrow \mathcal{P}_{zk}(\mathsf{pk}, (n, z_0, z_n), \Pi_n)$.
(3) The verifier $\mathcal{V}_{zk}$ takes the verification key vk, proof π, and (n, z_0, z_n) as inputs. It then either accepts the proof or rejects it.

If the zkSNARK is based on a Pedersen commitment scheme for vectors, then the proof size is $O(\log |F|)$ and the proof verification time is $O(|F|)$. Here $|F|$ is the number of R1CS constraints needed to express the computation of the step function F.

6 MPROVE-NOVA POR AND THE NON-COLLUSION PROTOCOL

MProve-Nova requires the specification of a Nova step function F that will prove ownership of an unspent Monero output (see definition in Section A.4) at each step and accumulate the coins in the output into a running sum. The necessity of the different components in F will become apparent when they are mapped to the requirements of any privacy-preserving Monero PoR protocol. These requirements are listed below:

R1. An exchange must only use outputs present on the Monero blockchain to contribute to its reserves.
R2. An exchange must only use outputs it owns to contribute to its reserves.
R3. An exchange must not use spent outputs to contribute to its reserves.
R4. An exchange must use each owned output at most once to contribute to its reserves.

6.1 Merkle Trees

To satisfy these requirements, we require one regular Merkle tree and two indexed Merkle trees [4]. The latter type of trees were introduced in [45] to generate efficient non-membership proofs inside a SNARK.

All three trees are constructed using the Poseidon hash function [21] to reduce the number of R1CS constraints used to express F. Let $H_{pos} : \{0,1\}^* \mapsto \mathbb{F}_s$ be the Poseidon hash function where $\mathbb{F}_s$ is the scalar field used to express the R1CS constraints.

(1) **Transaction Outputs Tree**: The *transaction outputs tree* (TXOT) is a regular Merkle tree. Let $H_p : \mathbb{G} \mapsto \mathbb{G}$ be the cryptographic hash function used to compute the key image of a one-time address $P = xG$ as $xH_p(P)$. Let $\|$ denote the string concatenation operator.
For a Monero block height bh, let $\mathcal{T}_{\mathsf{bh}} = \{(P_1, C_1), (P_2, C_2), \ldots\}$ be the set of all transaction outputs that have appeared in blocks upto height bh. TXOT is constructed using the leaves $\left\{H_{pos}\left(P \| C \| H_p(P)\right) \mid (P, C) \in \mathcal{T}_{\mathsf{bh}}\right\}$, sorted in the order of appearance of the one-time addresses (the P's) on the blockchain. We omit the dependence of TXOT on bh for simplicity. The motivation for choosing this particular leaf structure is described in Section 6.3.2.
(2) **Key Images Tree**: Let $\mathcal{I}_{\mathsf{bh}}$ be the set of all key images that have appeared on the Monero blockchain upto block height bh. The *key images tree* KIT is an indexed Merkle tree constructed using the leaves $\{H_{pos}(I) \mid I \in \mathcal{I}_{\mathsf{bh}}\}$, sorted in the order of appearance of the key images on the blockchain. Once again, we omit the dependence of KIT on bh for simplicity.
(3) **Double Spend Tree**: While the previous two trees are populated before the MProve-Nova protocol execution, the *double spend tree* DST is populated during the protocol execution. It is an indexed Merkle tree that is used to prevent the exchange from using an owned output more than once (hence satisfying requirement R4). The DST is initially empty and is later populated with leaves of the form $H_{pos}(x \| \mathsf{bh})$ where x is a private key. Let DST_{j-1} and DST_j denote the state of the double spend tree before and after the jth step, respectively.

6.2 Step Function Inputs and Outputs

Let n be the number of outputs the exchange will use to contribute to its reserves. For $j = 1, 2, \ldots, n$, the function F takes the following as **public inputs** in step j:

(i) The block height bh.
(ii) The root hash of TXOT.
(iii) The root hash of KIT.
(iv) The root hash of DST_{j-1}, the double spend tree before step j.
(v) A Pedersen commitment C^{res}_{j-1} to the reserves accumulated before step j. C^{res}_0 is set to G, which commits to the zero amount with blinding factor 1.

The **public outputs** of F in step j are the same as above, except for two differences.

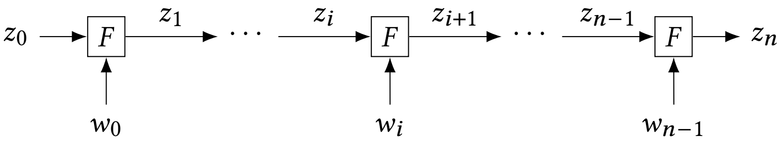

Figure 1: Incrementally Verifiable Computation

(i) The root hash of DST_{j-1} is replaced with the root hash of DST_j, the double spend tree *after* step j.

(ii) The Pedersen commitment C_{j-1}^{res} is replaced with C_j^{res}, the reserves accumulated *after* step j.

F also takes the following as **private inputs** in step j:

(i) An index i_j such that $1 \leq i_j \leq |\mathcal{T}_{\mathsf{bh}}|$.

(ii) The private key x_{i_j} corresponding to the one-time address P_{i_j}.

(iii) The values C_{i_j} and $H_p(P_{i_j})$ in the preimage of the leaf $H_{pos}\left(P_{i_j} \| C_{i_j} \| H_p(P_{i_j})\right)$ of TXOT at index i_j.

(iv) A Merkle path of a leaf in TXOT to prove the membership of the leaf $H_{pos}\left(P_{i_j} \| C_{i_j} \| H_p(P_{i_j})\right)$ in TXOT.

(v) A second Merkle path of a leaf in KIT to prove the non-membership of the Poseidon hash of the key image $I_{i_j} = x_{i_j} H_p(P_{i_j})$ in the KIT.

(vi) A third Merkle path of a leaf in DST_{j-1} to prove the non-membership of the leaf $H_{pos}(x_{i_j} \| \mathsf{bh})$ in DST_{j-1}.

(vii) A fourth Merkle path in DST_{j-1} to aid the insertion of the leaf $H_{pos}(x_{i_j} \| \mathsf{bh})$.

(viii) A Pedersen commitment C_{j-1}^{res} to the reserves accumulated before step j.

(ix) A random scalar r_j to blind the sum $C_{j-1}^{res} + C_{i_j}$.

6.3 Step Function Computation

For each $j \in \{1, 2, \ldots, n\}$, in step j function F executes the following substeps:

S1. Using the private input x_{i_j}, it computes the point $P_{i_j} = x_{i_j} G$. This computation proves knowledge of the private key corresponding to P_{i_j}.

S2. Using P_{i_j} and the private inputs C_{i_j} and $H_p(P_{i_j})$, it computes the hash $H_{pos}\left(P_{i_j} \| C_{i_j} \| H_p(P_{i_j})\right)$.

S3. Using the first private Merkle path input, it proves the membership of $H_{pos}(P_{i_j} \| C_{i_j} \| H_p(P_{i_j}))$ in TXOT.

S4. Using the private inputs x_{i_j} and $H_p(P_{i_j})$, it computes the key image $I_{i_j} = x_{i_j} H_p(P_{i_j})$ and its hash $H_{pos}(I_{i_j})$.

S5. Using the second private Merkle path input, it proves the non-membership of $H_{pos}(I_{i_j})$ in KIT.

S6. Using the private input x_{i_j}, it computes the hash $H_{pos}(x_{i_j} \| \mathsf{bh})$.

S7. Using the third private Merkle path input, it proves the non-membership of $H_{pos}(x_{i_j} \| \mathsf{bh})$ in DST_{j-1}.

S8. Using the fourth private Merkle path input, it inserts $H_{pos}(x_{i_j} \| \mathsf{bh})$ into DST_{j-1} to obtain the root hash of the new tree state DST_j.

S9. Using the private input r_j, it computes the point $C_j^{rand} = r_j G$.

S10. Using the private input C_{i_j}, public input C_{j-1}^{res}, and C_j^{rand}, it computes the updated Pedersen commitment $C_j^{res} = C_{j-1}^{res} + C_{i_j} + C_j^{rand}$.

The public outputs after the nth step contain C_n^{res} and the root hash of DST_n.

- The exchange will claim that C_n^{res} is a Pedersen commmitment to its total reserves. The exchange can generate a range proof on the point $C_n^{res} - tH$ to prove that its reserves exceed a threshold t.
- The root hash of DST_n will be used as input to the non-collusion protocol.

6.3.1 Satisfying Monero PoR requirements. Let us revisit the requirements listed at the beginning of this section.

(i) Steps S2 and S3 prove that C_{i_j} is a Pedersen commitment to coins tied to the one-time address P_{i_j} present on the Monero blockchain. Steps S9 and S10 ensure that only such C_{i_j}'s contribute to the total reserves of the exchange. Thus the exchange can only use outputs from the Monero blockchain to contribute to its reserves (requirement R1).

(ii) Step S1 proves that the exchange knows the private key x_{i_j} corresponding to P_{i_j}. Thus requirement R2 is satisfied, i.e. the exchange only uses outputs it owns to contribute to its reserves.

(iii) Steps S4 and S5 prove that the exchange can only use unspent outputs to contribute to its reserves, thus satisfying requirement R3. If an exchange were to use a spent output, the corresponding key image would be present in KIT and the non-membership proof in S5 would fail.

(iv) Steps S6, S7, S8 prove that the coins tied to the one-time address P_{i_j} are used at most once to contribute to the exchange's reserves, thus satisfying requirement R4.
  - Suppose an exchange attempts to use the same output twice, in steps j_1 and j_2 where $j_2 > j_1$.
  - It would have to repeat the private key x to satisfy the membership proof in step S3.
  - The leaf $H_{pos}(x \| \mathsf{bh})$ would be inserted into the double spend tree in substep S8 of step j_1 to get DST_{j_1}.
  - Then the non-membership proof of $H_{pos}(x \| \mathsf{bh})$ in DST_{j_2-1} would fail in substep S7 of the later step j_2.

6.3.2 Motivating the Step Function Structure. In this subsection, we justify our choices for the various data structures and computations used to specify the step function.

(i) We chose the TXOT to have leaves of the form $H_{pos}\left(P_{i_j} \| C_{i_j} \| H_p(P_{i_j})\right)$. It may not be clear why we chose to concatenate C_{i_j} and $H_p(P_{i_j})$ with P_{i_j}.

- The reason for including C_{i_j} in the leaf is to enforce a *copy constraint*. We want to ensure that only the commitment C_{i_j} tied to P_{i_j} contributes to the reserves in step j. It is possible for users to know the amounts and blinding factors of commitments tied to one-time addresses they *do not own* (see Section A.3). We want to avoid situations where commitments not owned by the exchange are used to generate its reserves.
- The point $H_p(P_{i_j})$ is included in the leaf to reduce the number of R1CS constraints needed to specify the step function F. If it were not included, then the exchange would have to calculate the hash of P_{i_j} in the step function. This hash involves the Keccak hash function [8], which requires a large number of R1CS constraints.

(ii) We chose to insert leaves of the form $H_{pos}(x_{i_j}||\mathsf{bh})$ into the DST to prevent an exchange from using an owned output more than once. Such behavior could also be possibly prevented by inserting any one of $H_{pos}(x_{i_j})$, $H_{pos}(P_{i_j})$ or $H_{pos}(P_{i_j}||\mathsf{bh})$ into the DST.

- The problem with using $H(P_{i_j})$ or $H_{pos}(P_{i_j}||\mathsf{bh})$ as leaves to the DST is that an adversary could try combinations of outputs to reconstruct the set of owned outputs from the public root hash of the DST. This attack is feasible for small n as the number of outputs on the Monero blockchain is less than 100 million [48, 51].
- The problem with using $H(x_{i_j})$ as leaves to the DST is that the public root hash of the DST would remain the same if the exchange uses the *same set of outputs* to generate its reserves at *different* block heights. This leaks information about the asset strategy of the exchange.

(iii) In substep S10 of step j, the Pedersen commitment to the accumulated reserves C_j^{res} is blinded using a random point C_j^{rand}. The point C_j^{rand} is chosen to have the form r_jG to ensure that the amount being contributed by C_{i_j} is unchanged (see Section A.3). If this blinding point were not present and the accumulated reserves were simply calculated as $C_j^{res} = C_{j-1}^{res} + C_{i_j}$, then an adversary could try combinations of commitments to reconstruct the set of owned outputs from the public output C_n^{res}.

6.4 Protocol for Proving Non-Collusion

To generate a proof of non-collusion, we assume that the values $H_{pos}(x_{i_j}||\mathsf{bh})$ which were inserted into the DST in step j of the PoR protocol are made public by the exchange. We argue in Section B.2, that this does not affect the exchange's privacy.

If all the exchanges were to make the leaves of their respective DSTs public, then customers can check for non-collusion by checking that these sets are all pairwise disjoint. But the customers would need to download all the leaves of the DST and the time taken to check non-collusion would be proportional to the size of leaf sets.

Instead, we propose a protocol for proving non-collusion between a pair of exchanges that leverages Nova to achieve constant verification time and reduces the download requirements to just the root hash of the DST and a constant-sized Nova proof. For this protocol to work, the exchanges need to generate their respective MProve-Nova proofs (which contains their DSTs) at the same block height bh.

Suppose two exchanges Ex1 and Ex2, that have generated MProve-Nova proofs of reserves at the same blockheight bh, want to prove that they have not colluded. They want to prove that they have not used any common outputs to contribute to their respective reserves.

(i) Let n_1 and n_2 be the number of owned outputs of Ex1 and Ex2, respectively.
(ii) Let $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$ and $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$ be the double spend trees obtained by the exchanges at the end of their MProve-Nova proof procedures.
(iii) Let $v_1^{(1)}, v_2^{(1)}, \ldots, v_{n_1}^{(1)}$ be the leaves of $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$ and $v_1^{(2)}, v_2^{(2)}, \ldots, v_{n_2}^{(2)}$ be the leaves of $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$.
(iv) Note that $v_j^{(1)} = H_{pos}\left(x_j^{(1)}||\mathsf{bh}\right)$ for $j = 1, 2, \ldots, n_1$ and $v_j^{(2)} = H_{pos}\left(x_j^{(2)}||\mathsf{bh}\right)$ for $j = 1, 2, \ldots, n_2$, where the $x_j^{(1)}$s and $x_j^{(2)}$s are the private keys owned by Ex1 and Ex2, respectively.

The non-collusion protocol requires an indexed Merkle tree called the *output inclusion tree* OIT that serves a purpose similar to the double spend tree DST in the proof of reserves protocol. The OIT is initially empty and is later populated with leaves of the DST of one of the exchanges. Let OIT_{j-1} and OIT_j denote the state of the output inclusion tree before and after the jth step, respectively.

6.4.1 Step Function Inputs and Outputs.

Ex2 will share the leaves $v_1^{(2)}, v_2^{(2)}, \ldots, v_{n_2}^{(2)}$ of its double spend tree with Ex1. The Nova step function F_{nc} for the non-collusion protocol will be **run by** Ex1 for n_1 steps. It takes the following as **public inputs** in step j:

(i) The root hashes of the double spend trees $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$ and $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$.
(ii) The root hash of OIT_{j-1}, the output inclusion tree *before* step j.

The **public outputs** of F_{nc} in step j are the following:

(i) The root hashes of the double spend trees $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$ and $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$.
(ii) The root hash of OIT_j, the output inclusion tree *after* step j.

The step function F_{nc} also takes the following as **private inputs** in step j:

(i) A Merkle path to prove the membership of the leaf $v_j^{(1)}$ in $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$.
(ii) A second Merkle path to prove non-membership of $v_j^{(1)}$ in $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$.
(iii) A third Merkle path to prove the non-membership of the leaf $v_j^{(1)}$ in OIT_{j-1}.
(iv) A fourth Merkle path in OIT_{j-1} to aid the insertion of the leaf $v_j^{(1)}$ into it.

6.4.2 Step Function Computation.

For $j \in \{1, .., n_1\}$, in step j, function F_{nc} executes the following substeps:

S1. Using the first private Merkle path input, it proves the membership of $v_j^{(1)}$ in $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$.

S2. Using the second private Merkle path input, it proves the non-membership of $v_j^{(1)}$ in $\mathsf{DST}_{n_2}^{\mathsf{Ex2}}$.

S3. Using the third private Merkle path input, it proves the non-membership of $v_j^{(1)}$ in OIT_{j-1}.

S4. Using the fourth private Merkle path input, it inserts $v_j^{(1)}$ into OIT_{j-1} to obtain the root hash of the new tree state OIT_j.

After step n_1, the root hash of OIT_{n_1} will equal the root hash of $\mathsf{DST}_{n_1}^{\mathsf{Ex1}}$.

6.4.3 Motivating the Step Function Structure. The motivation for the different substeps in F_{nc} is as follows:

(i) Steps S1 and S2 prove that an output used by Ex1 is not used by Ex2.

(ii) Steps S3 and S4 prove that the output $v_j^{(1)}$ is distinct from $v_1^{(1)}, v_2^{(1)}, \ldots, v_{j-1}^{(1)}$.

7 SECURITY PROPERTIES OF MPROVE-NOVA

The MProve-Nova PoR protocol has the following security properties:

(1) *Inflation Resistance*: The inflation resistance property requires that a computationally bounded exchange cannot create a commitment to an amount which is greater than its total reserves.

(2) *Exchange Privacy*: The exchange privacy property prevents a computationally bounded adversary from identifying the Monero outputs belonging to the exchange.

(3) *Collusion Resistant*: The output of the MProve-Nova PoR can be used as input to a protocol for proving non-collusion. The latter protocol cannot be violated by computationally bounded exchanges, provided that they both generated their proofs of reserves at the same block height.

These security properties and their proofs are further discussed in Appendix B.

8 IMPLEMENTATION AND PERFORMANCE

We have implemented the MProve-Nova PoR protocol and the protocol for proving non-collusion using our modified implementation of Nova [44]. The implementation uses Pasta curves [52], a 2-cycle of elliptic curves and the Posiedon hash function [21, 28]. A detailed explanation of Nova using 2-cycle of elliptic curves and a security fix in the implementation was given by Nguyen *et al.* [38]. The Nova implementation [44] accepts the step computation F, as a bellpepper gadget [29]. The bellpepper gadgets required to implement MProve-Nova were described in Section 3.

In this section, we present the performance of the MProve-Nova PoR protocols. All simulations were run on a 64 core 2.30GHz Intel Xeon Gold 6314U CPU with access to 125GB RAM. Table 1 shows comparison between MProve-Nova, MProve+, and MProve for PoR generation/verification. The open source implementations of MProve+ [6] and MProve [5] were used to perform the simulations. Our implementation of MProve-Nova is available on anonymous Github [37].

The computation times and proof sizes of MProve+ and MProve increase linearly in the size of the anonymity set. All simulations for MProve+ and MProve were run for an anonymity set of 45,000 one-time addresses, whereas for MProve-Nova the anonymity set is the set of all one-time addresses on the Monero blockchain. Thus the comparison is unfair towards MProve-Nova since for an anonymity set of all the one-time addresses, MProve+ and MProve will be impractical. However, despite this, MProve-Nova gives practical results and performs better in terms of verification times and proof sizes.

For MProve-Nova, the proving time is linear in the number of owned one-time addresses, while the proof verification times and proof sizes are constant. The proving time for 10,000 owned addresses is about 7 hours. The verification time and proof size are constant at about 4.5 s and 27 KB, respectively. The performance results are practical since an exchange with up to 10,000 owned addresses can generate proofs every 12 Hrs and an exchange with up to 20,000 owned addresses can generate proofs every 24 Hrs.

MProve-Nova has smaller proof sizes and faster verification time as compared to both MProve+ and MProve. MProve-Nova has higher proving time compared to MProve but it provides better privacy to the exchange which is not the case with MProve.

Table 2 shows the performance of the non-collusion protocol. In all cases, the proof size is 23 KB and proof verification time is about 219 ms.

9 CONCLUSION

We described MProve-Nova, the first privacy-preserving proof of reserves for Monero which also enables proofs of non-collusion between exchanges. We compared MProve-Nova with MProve+ and MProve to show that our protocol has practical computation times and proof size along with better privacy.

For MProve-Nova, the proving time for 10,000 owned addresses is about 7 hours. It achieves verification times of about 4.5 seconds and proof sizes of 27 KB, irrespective of the size of the anonymity set. Our goal with the performance evaluation is to provide an upperbound on the computation times since they can be further reduced using GPU acceleration or using non-uniform IVC schemes like SuperNova [25].

REFERENCES

[1] 0xPARC. 2021. circom-ecdsa: Implementation of ECDSA operations in circom. https://github.com/0xPARC/circom-ecdsa

[2] Zachary Amsden and et al. 2020. The Libra Blockchain. https://diem-developers-components.netlify.app/papers/the-diem-blockchain/2020-05-26.pdf

[3] Sebastian Angel. 2023. Nova Implementation with Zero-knowledge. https://github.com/sga001/Nova

[4] Aztec. 2023. Indexed Merkle Tree. https://docs.aztec.network/aztec/protocol/trees/indexed-merkle-tree

[5] Suyash Bagad. 2020. Implementation of MProve. https://github.com/suyash67/MProve-Ristretto

[6] Suyash Bagad. 2021. Implementation of MProve+. https://github.com/suyash67/MProvePlus-Ristretto

[7] Foteini Baldimtsi, Panagiotis Chatzigiannis, S. Gordon, Phi Le, and Daniel McVicker. 2022. gOTzilla: Efficient Disjunctive Zero-Knowledge Proofs from MPC in the Head, with Application to Proofs of Assets in Cryptocurrencies. *Proceedings on Privacy Enhancing Technologies* 2022 (10 2022), 229–249. https://doi.org/10.56553/popets-2022-0107

Table 1: Performance comparison of PoR protocols. $n = |\mathcal{P}_{own}|$, PT denotes proving time, VT denotes verification time and PS denotes proof size with units in parentheses. Values with a * are estimated values due to simulation running out of memory.

	MProve-Nova			*MProve+*			*MProve*		
n	PT(Hrs)	VT(s)	PS(KB)	PT(Hrs)	VT(s)	PS(KB)	PT(s)	VT(s)	PS(MB)
1,000	0.66	4.54	27.37	2.49	959.05	162.62	99.32	17.36	37.44
3,000	1.99	4.50	27.36	7.46*	2887.54*	482.81*	100.19	17.33	37.44
5,000	3.36	4.53	27.37	12.44*	4817.22*	803.02*	100.21	17.37	37.44
7,000	4.75	4.57	27.36	17.41*	6746.34*	1123.2*	100.27	17.48	37.44
10,000	6.94	4.48	27.36	24.88*	9640.44*	1603.5*	100.19	17.34	37.44

Table 2: Performance of Proof of Non-Collusion. n_1 denotes the number of values inserted by exchange Ex1 in its double spend tree $\mathsf{DST}^{\mathsf{Ex1}}$

n_1	Proving Time	Verification Time	Proof Size
1,000	359.66 s	209.30 ms	23.18 KB
3,000	1165.06 s	222.85 ms	23.18 KB
5,000	2373.40 s	219.54 ms	23.18 KB
7,000	3944.93 s	231.37 ms	23.18 KB
10,000	6407.71 s	218.58 ms	23.18 KB

[8] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. 2014. The making of KECCAK. *Cryptologia* 38, 1 (2014), 26–60. https://doi.org/10.1080/01611194.2013.856818

[9] Nalin Bhardwaj. 2022. Nova Scotia: Middleware to compile Circom circuits to Nova prover. https://github.com/nalinbhardwaj/Nova-Scotia

[10] Gautam Botrel, Thomas Piellard, Youssef El Housni, Ivo Kubjas, and Arya Tabaie. 2022. *ConsenSys/gnark: v0.7.0.* https://doi.org/10.5281/zenodo.5819104

[11] B. Bünz, J. Bootle, D. Boneh, A. Poelstra, P. Wuille, and G. Maxwell. 2018. Bulletproofs: Short Proofs for Confidential Transactions and More. In *2018 IEEE Symposium on Security and Privacy (IEEE S&P).* 315–334. https://doi.org/10.1109/SP.2018.00020

[12] ChainSec. 2023. The Complete List of Crypto Exchange Hacks - ChainSec — chainsec.io. https://chainsec.io/exchange-hacks/

[13] Konstantinos Chalkias, Panagiotis Chatzigiannis, and Yan Ji. 2022. Broken Proofs of Solvency in Blockchain Custodial Wallets and Exchanges. Cryptology ePrint Archive, Paper 2022/043. https://eprint.iacr.org/2022/043

[14] Konstantinos Chalkias, Kevin Lewi, Payman Mohassel, and Valeria Nikolaenko. 2020. Distributed Auditing Proofs of Liabilities. Cryptology ePrint Archive, Paper 2020/468. https://eprint.iacr.org/2020/468

[15] Panagiotis Chatzigiannis, Foteini Baldimtsi, and Konstantinos Chalkias. 2021. SoK: Auditability and Accountability in Distributed Payment Systems. Cryptology ePrint Archive, Paper 2021/239. https://eprint.iacr.org/2021/239

[16] Panagiotis Chatzigiannis and Konstantinos Chalkias. 2021. Proof of Assets in the Diem Blockchain. In *Applied Cryptography and Network Security Workshops: ACNS 2021 Satellite Workshops, AIBlock, AIHWS, AIoTS, CIMSS, Cloud S&P, SCI, SecMT, and SiMLA, Kamakura, Japan, June 21–24, 2021, Proceedings* (Kamakura, Japan). Springer-Verlag, Berlin, Heidelberg, 27–41. https://doi.org/10.1007/978-3-030-81645-2_3

[17] Consensys. 2020. Implementation of gnark emulated package. https://github.com/Consensys/gnark/tree/master/std/math/emulated

[18] Gaby G. Dagher, Benedikt Bünz, Joseph Bonneau, Jeremy Clark, and Dan Boneh. 2015. Provisions: Privacy-Preserving Proofs of Solvency for Bitcoin Exchanges. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (Denver, Colorado, USA) *(CCS '15).* Association for Computing Machinery, New York, NY, USA, 720–731. https://doi.org/10.1145/2810103.2813674

[19] Arijit Dutta, Suyash Bagad, and Saravanan Vijayakumaran. 2021. MProve+: Privacy Enhancing Proof of Reserves Protocol for Monero. *IEEE Transactions on Information Forensics and Security* 16 (2021), 3900–3915. https://doi.org/10.1109/TIFS.2021.3088035

[20] Arijit Dutta and Saravanan Vijayakumaran. 2019. MProve: A Proof of Reserves Protocol for Monero Exchanges. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW).* 330–339. https://doi.org/10.1109/EuroSPW.2019.00043

[21] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. 2021. Poseidon: A New Hash Function for Zero-Knowledge Proof Systems. In *30th USENIX Security Symposium (USENIX Security 21).* USENIX Association, 519–535. https://www.usenix.org/conference/usenixsecurity21/presentation/grassi

[22] iden3. 2024. Circom circuit compiler. https://github.com/iden3/circom

[23] Yan Ji and Konstantinos Chalkias. 2021. Generalized Proof of Liabilities. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security* (Virtual Event, Republic of Korea) *(CCS '21).* Association for Computing Machinery, New York, NY, USA, 3465–3486. https://doi.org/10.1145/3460120.3484802

[24] Koe, Kurt M. Alonso, and Sarang Noether. 2020. Zero to Monero: Second Edition. https://www.getmonero.org/library/Zero-to-Monero-2-0-0.pdf

[25] Abhiram Kothapalli and Srinath Setty. 2022. SuperNova: Proving universal machine executions without universal circuits. Cryptology ePrint Archive, Paper 2022/1758. https://eprint.iacr.org/2022/1758

[26] Abhiram Kothapalli, Srinath Setty, and Ioanna Tzialla. 2022. Nova: Recursive Zero-Knowledge Arguments from Folding Schemes. In *Advances in Cryptology – CRYPTO 2022: 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15–18, 2022, Proceedings, Part IV* (Santa Barbara, CA, USA). Springer-Verlag, Berlin, Heidelberg, 359–388. https://doi.org/10.1007/978-3-031-15985-5_13

[27] Amrit Kumar, Clément Fischer, Shruti Tople, and Prateek Saxena. 2017. A Traceability Analysis of Monero's Blockchain. In *Computer Security – ESORICS 2017*, Simon N. Foley, Dieter Gollmann, and Einar Snekkenes (Eds.). Springer International Publishing, Cham, 153–173. https://doi.org/10.1007/978-3-319-66399-9_9

[28] Lurk Lab. 2020. neptune : Implementation of the Poseidon hash function. https://github.com/lurk-lab/neptune

[29] Lurk Lab. 2023. bellpepper : Rust Library for R1CS circuits. https://github.com/lurk-lab/bellpepper

[30] Russell W. F. Lai, Viktoria Ronge, Tim Ruffing, Dominique Schröder, Sri Aravinda Krishnan Thyagarajan, and Jiafan Wang. 2019. Omniring: Scaling Private Payments Without Trusted Setup. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom) *(CCS '19).* Association for Computing Machinery, New York, NY, USA, 31–48. https://doi.org/10.1145/3319535.3345655

[31] Joseph K. Liu, Victor K. Wei, and Duncan S. Wong. 2004. Linkable Spontaneous Anonymous Group Signature for Ad Hoc Groups. In *Information Security and Privacy*, Huaxiong Wang, Josef Pieprzyk, and Vijay Varadharajan (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 325–335. https://doi.org/10.1007/978-3-540-27800-9_28

[32] Microsoft. 2021. Nova Implementation. https://github.com/microsoft/Nova

[33] Microsoft. 2023. Zero-knowledge implementation gap in Nova. https://github.com/microsoft/Nova/issues/174

[34] Monero. 2024. The Monero Project. https://www.getmonero.org/

[35] MoneroSchedule 2020. Monero Scheduled Software Upgrades. https://github.com/monero-project/monero/#scheduled-software-upgrades Last Accessed: August 13, 2023.

[36] Malte Möser, Kyle Soska, Ethan Heilman, Kevin Lee, Henry Heffan, Shashvat Srivastava, Kyle Hogan, Jason Hennessey, Andrew Miller, Arvind Narayanan, and Nicolas Christin. 2018. An Empirical Analysis of Traceability in the Monero Blockchain. *Proceedings on Privacy Enhancing Technologies* 2018, 3 (2018), 143–163. https://doi.org/10.1515/popets-2018-0025

[37] MProve Nova. 2024. Implementation of MProve-Nova. https://anonymous.4open.science/r/5t384rtcbf57fkbvksdncoir893457022f674r3658h32y8cxny87/README.md

[38] Wilson Nguyen, Dan Boneh, and Srinath Setty. 2023. Revisiting the Nova Proof System on a Cycle of Curves. Cryptology ePrint Archive, Paper 2023/969. https://eprint.iacr.org/2023/969

[39] Shen Noether, Adam Mackenzie, and The Lab. 2016. Ring Confidential Transactions. *Ledger* 1 (12 2016), 1–18. https://doi.org/10.5195/LEDGER.2016.34

[40] Torben Pryds Pedersen. 1992. Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In *Advances in Cryptology — CRYPTO '91*. Springer, 129–140. https://doi.org/10.1007/3-540-46766-1_9

[41] Elements Project. 2018. Proof-of-Reserves tool for Bitcoin. https://github.com/ElementsProject/reserves

[42] Steven Roose. 2019. Standardizing Bitcoin Proof of Reserves. https://blog.blockstream.com/en-standardizing-bitcoin-proof-of-reserves/

[43] Stoffu Noether. 2018. Reserve Proof Pull Request. https://github.com/monero-project/monero/pull/3027

[44] Varun Thakore. 2024. Nova Implementation with Zero-knowledge. https://github.com/varunthakore/Nova

[45] Ioanna Tzialla, Abhiram Kothapalli, Bryan Parno, and Srinath Setty. 2022. Transparency Dictionaries with Succinct Proofs of Correct Operation. *ISOC Conference on Network and Distributed System Security (NDSS)* (2022). https://doi.org/10.14722/ndss.2022.23143

[46] Paul Valiant. 2008. Incrementally Verifiable Computation or Proofs of Knowledge Imply Time/Space Efficiency. In *Proceedings of the 5th Conference on Theory of Cryptography* (New York, USA) *(TCC'08)*. Springer-Verlag, Berlin, Heidelberg, 1–18. https://doi.org/10.1007/978-3-540-78524-8_1

[47] Nicolas van Saberhagen. 2013. CryptoNote v 2.0. https://bytecoin.org/old/whitepaper.pdf

[48] Saravanan Vijayakumaran. 2021. Analysis of CryptoNote Transaction Graphs using the Dulmage-Mendelsohn Decomposition. Cryptology ePrint Archive, Paper 2021/760. https://eprint.iacr.org/2021/760

[49] Wikipedia. 2023. FTX — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/FTX

[50] Wikipedia. 2023. Mt. Gox — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Mt._Gox

[51] Zuoxia Yu, Man Ho Au, Jiangshan Yu, Rupeng Yang, Qiuliang Xu, and Wang Fat Lau. 2019. New Empirical Traceability Analysis of CryptoNote-Style Blockchains. In *Financial Cryptography and Data Security*. 133–149. https://doi.org/10.1007/978-3-030-32101-7_9

[52] Zcash. 2020. Pasta curves. https://github.com/zcash/pasta

[53] zkcrypto. 2015. bellman : Rust Library for R1CS circuits. https://github.com/zkcrypto/bellman

A OVERVIEW OF MONERO

Monero [34] is the most popular instantiation of the CryptoNote protocol [47], with additional privacy and efficiency improvements. In Monero transactions, receiver identities are hidden using *one-time addresses*, sender identities are obfuscated using *linkable ring signatures*, and the number of coins being transferred is hidden using *Pedersen commitments* [40].

A.1 One-Time Addresses

Monero public keys are points in the prime order subgroup of the Twisted Edwards elliptic curve Ed25519 [24]. Let $\mathbb{G}$ denote this subgroup whose order is a 253-bit prime l. Monero private keys are integers in the set $\mathbb{Z}_l^+ = \{1, 2, ..., l-1\}$. For the basepoint $G \in \mathbb{G}$, the public key $P \in \mathbb{G}$ corresponding to a private key $x \in \mathbb{Z}_l^+$ is denoted by $P = xG$. We will use additive notation for scalar multiplication throughout this paper.

Suppose Alice wants to send some Monero coins to Bob.

(1) Bob shares a public key pair $(B_{vk}, B_{sk}) \in \mathbb{G}^2$ with Alice. The subscripts vk and sk are abbreviations of *view key* and *spend key*. Let $(b_{vk}, b_{sk}) \in \mathbb{Z}_l^+ \times \mathbb{Z}_l^+$ denote the corresponding private key pair.

(2) She signs a transaction transferring coins she owns to Bob. This transaction will contain a *one-time address* P that will be controlled by Bob and a random point R.

(3) Alice creates the one-time address P as follows:

  (i) She chooses a random scalar $r \in \mathbb{Z}_l^+$.

  (ii) She computes the one-time address as

$$P = H(rB_{vk} \| o_{index})G + B_{sk}$$

  where $\|$ denotes concatenation, $H : \{0,1\}^* \mapsto \mathbb{Z}_l^+$ is a cryptographic hash function, and o_{index} is the index of the new *output* (to be defined later) in the transaction.[2] Note that the private key corresponding to P is known *only* to Bob as it equals $x = H(rB_{vk} \| o_{index})G + b_{sk}$ where $B_{sk} = b_{sk}G$.

(4) Alice computes the random point R as rG.

(5) Alice broadcasts the transaction containing (P, R), which will be eventually included in a Monero block by miners.

(6) Bob identifies transactions transferring coins to him as follows:

  (i) For every new Monero block, Bob reads the point pairs (P, R) in all the transactions.

  (ii) He computes the point $P' = H(b_{vk}R \| o_{index})G + B_{sk}$.

  (iii) If $P' = P$, Bob concludes that the transaction is sending coins to him.

(7) Bob adds P to the list of one-time addresses owned by him.

Note that Bob will always be able to identify transactions meant for him as $rB_{vk} = rb_{vk}G = b_{vk}R$. This is nothing but a Diffie-Hellman shared secret between public keys R and B_{vk}.

One-time addresses generated using Bob's public key pair cannot be linked to his key pair as long as the decisional Diffie-Hellman (DDH) problem remains hard in Ed25519. In this way, Monero hides the receiver's identity in a transaction.

A.2 Linkable Ring Signatures

Monero uses linkable ring signatures [31, 39] to obfuscate sender identities, while preventing double spending. Given a list of public keys, a ring signature allows a signer to prove that he knows the private key of one public key from the list *without* revealing which one. A linkable ring signature allows an observer to link multiple ring signatures generated using the same private key.

Suppose Bob wants to spend the coins tied to a one-time address P he owns, i.e. he knows $x \in \mathbb{Z}_l^+$ such that $P = xG$. This one-time address is already present on the Monero blockchain. He proceeds as follows:

(1) For a protocol-specified ring size n, Bob randomly samples $n-1$ one-time addresses $P_1, P_2, \ldots, P_{n-1}$ (all distinct from P) from the blockchain. These are called *decoy* addresses.

(2) Bob signs the spending transaction using a linkable ring signature on the set of one-time addresses $\mathcal{P} = \{P_1, P_2, \ldots, P_{n-1}, P\}$. This set is sorted in chronological order (oldest address first) to prevent the ordering of the keys in $\mathcal{P}$ from leaking the identity of P.

(3) Bob includes the linkable ring signature in the transaction he broadcasts to the Monero P2P network.

Hiding the identity of the spending key opens up the possibility of double spending. To prevent this, the linkable ring signature contains an Ed25519 point called the *key image*, defined as $I = xH_p(P)$ where $H_p : \mathbb{G} \mapsto \mathbb{G}$ is a point-valued cryptographic hash function.

Two linkable ring signatures spending from the same one-time address will have identical key images. The Monero blockchain

[2] The output index is included to allow the creation of distinct one-time addresses from the same public key pair in the same transaction.

maintains the set $\mathcal{I}$ of key images that have appeared in past transactions. If the coins tied to a one-time address P have already been spent, then its key image I will already be in $\mathcal{I}$. Monero block miners will reject transactions whose linkable ring signatures have key images from $\mathcal{I}$.

At the same time, revealing the key image of a one-time address does not leak information about the latter as long as the DDH problem remains hard in Ed25519. To see this, let $H_p(P) = yG$ for some unknown $y \in \mathbb{Z}_l^+$. Then $I = xH_p(P) = xyG$ is the Diffie-Hellman function of $P = xG$ and $H_p(P)$.

A.3 Pedersen Commitments to Amounts

In the CryptoNote protocol specification, the number of coins tied to a one-time address was public. To create a ring signature spending from an address, the spender could only sample from other addresses containing the same amount.

To improve privacy, Monero introduced the use of Pedersen commitments [40] to hide the number of coins tied to a one-time address. The Pedersen commitment to an amount $a \in \{0, 1, 2, ..., 2^{64} - 1\}$ is given by

$$C(y, a) = yG + aH,$$

where $y \in \mathbb{Z}_l^+$ is a randomly chosen *blinding factor* and $H \in \mathbb{G}$ is a curve point whose discrete logarithm with respect to the base point G is unknown. Such commitments are perfectly hiding and computationally binding.

When Alice wants to transfer some of her coins to Bob, she creates a Pedersen commitment $C(a, y)$ in addition to the one-time address P. Alice includes $a' = a \oplus H_K(H_K(rB_{vk}))$ and $y' = y \oplus H_K(rB_{vk})$ in the transaction, where $\oplus$ is bitwise XOR and H_K is the Keccak hash function. Bob can recover a, y from a', y' using his private view key b_{vk} and the random point R in the transaction.

A.4 Monero Outputs

The destination of coin transfers in a Monero transaction is called an *output*. A transaction can have multiple outputs. Each output is characterized by a pair $(P, C) \in \mathbb{G}^2$, where P is a one-time address and C is a Pedersen commitment to the number of coins stored in the output.

Outputs containing Pedersen commitments are created in a Monero transaction type called *ring confidential transaction (RingCT)* [39]. Monero made the RingCT type of transactions mandatory in September 2017 [35].

In our design of MProve-Nova, we only consider RingCT outputs. In case an exchange owns a non-RingCT output, they can use a Pedersen commitment with zero blinding factor, i.e. a commitment of the form $C(0, a) = aH$, to represent the output.

B SECURITY PROPERTIES

In this section, we discuss the security properties of the MProve-Nova and non-collusion protocols. We model computationally bounded entities as probabilistic polynomial-time (PPT) algorithms.

B.1 Inflation Resistance

The inflation resistance property prevents a PPT exchange from creating a commitment to an amount which is greater than its total reserves. At each step j of the protocol, the exchange is allowed to accumulate the commitment C_{i_j} into the total commitment C_j^{res} if and only if

(i) it knows the private key corresponding to P_{i_j},
(ii) the output corresponding to (P_{i_j}, C_{i_j}) is unspent, and
(iii) the output corresponding to (P_{i_j}, C_{i_j}) has not been used by the exchange before.

These properties are enforced by the circuit of Nova step computation described in Section 6.

Suppose an exchange executes the MProve-Nova protocol for n steps where it uses the output (P_{i_j}, C_{i_j}) in the jth step. Let a_j be the amount committed by C_{i_j}. The following theorem shows that a PPT exchange can only use MProve-Nova to output a commitment to an amount which is a sum of the reserves it owns.

Theorem B.1. *Suppose an exchange can create a MProve-Nova proof of reserves with final commitment $C_n^{res} = C(y, a)$ such that*

(i) it knows the blinding factor $y \in \mathbb{Z}_l^+$ and an amount $a \in \{0, 1, \ldots, 2^{64} - 1\}$
(ii) the amount a is not of the form $\sum_{j=1}^n a_j$ and
(iii) the proof is accepted by the MProve-Nova verifier.

Then the exchange can calculate the discrete logarithm of H with respect to G with overwhelming probability.

Proof. As per the Nova step computation described in Section 6.2,

$$C_n^{res} = C_0^{res} + \sum_{j=1}^{n} \left(C_{i_j} + C_j^{rand} \right) = G + \sum_{j=1}^{n} \left(y_{i_j} G + a_j H + r_j G \right),$$

where y_{i_j} is the blinding factor in C_{i_j}. Now substitute $C_n^{res} = C(y, a) = yG + aH$

$$yG + aH = G + \sum_{j=1}^{n} \left(y_{i_j} G + a_j H + r_j G \right)$$

$$\left(a - \sum_{j=1}^{n} a_j \right) H = \left(1 - y + \sum_{j=1}^{n} y_j + r_j \right) G$$

The exchange can calculate the discrete logarithm of H with respect to G as

$$\left(a - \sum_{j=1}^{n} a_j \right)^{-1} \left(1 - y + \sum_{j=1}^{n} y_j + r_j \right)$$

The multiplicative inverse of $\left(a - \sum_{j=1}^n a_j \right)$ exists because it is a non-zero element in the prime field $\mathbb{F}_l$. This follows from the assumption that $a \neq \sum_{j=1}^n a_j$. □

B.2 Exchange Privacy

We will separately describe exchange privacy for the MProve-Nova PoR protocol and the non-collusion protocol.

B.2.1 Proof of Reserves. The exchange privacy property prevents a PPT adversary from identifying the Monero outputs which the exchange used to generate the proof of reserves. In the PoR protocol, the exchange produces an IVC proof Π_n that attests to the correct execution of steps 1, ..., n. Then the exchange uses the prover $\mathcal{P}_{zk}$ to produce a proof π to show that it knows a valid IVC proof Π_n

for the statement (z_0, z_n), where z_0 is the initial public input and z_n is the public output after the nth step. The proof π is submitted to the verifier or customer along with the statement (z_0, z_n). As described in Section 5, the proof π is zero-knowledge and it reveals nothing about the exchange's private inputs to the protocol.

The initial input z_0 is independent of the exchange's private inputs and thus reveals nothing about the exchange's private inputs. The final output of the protocol z_n depends on the exchange's private inputs as

$$z_n = F\left(F\left(\ldots F\left(F\left(F\left(z_0, w_0\right), w_1\right), w_2\right), \ldots\right), w_{n-1}\right).$$

We need to show that the output z_n does not reveal any information about the exchange's private input.

We consider the scenario when an exchange publishes $f(\lambda)$ number of PoR outputs, where $f(\lambda)$ denotes a polynomial of the security parameter λ. Each of the PoR outputs can be for different number of Monero outputs owned by the exchange. Let us denote the number of steps in kth output by n_k and the corresponding output of the protocol by z_{n_k}. We denote the $f(\lambda)$ outputs of the protocol as $\mathcal{Z}_{\text{actual}} = \{z_{n_k} \mid k = 1, \ldots, f(\lambda)\}$.

All the $f(\lambda)$ actual outputs of the protocol correspond to a different block heights bh of the Monero blockchain. If the exchange uses the same block height for different outputs then the set of private keys used to generate proof of reserves should not be same.

We want to show that $\mathcal{Z}_{\text{actual}}$ does not reveal any information of the exchange's private inputs. For this, the exchange constructs simulated outputs. For the kth actual output z_{n_k},

$$z_{n_k} = [\mathsf{bh}_k,\ \mathsf{KIT}_k.root,\ \mathsf{TXOT}_k.root,\ \mathsf{DST}_k.root,\ C^{res}_{n_k}]$$

the simulated output $\hat{z}_{n_k}$ is constructed as follows.

- The block height bh_k, key image tree root $\mathsf{KIT}_k.root$ and $\mathsf{TXOT}_k.root$ are set to be same as in z_{n_k}, since they are public data available on the Monero blockchain.
- The exchange samples n_k number of uniformly random private keys from $\mathbb{Z}_l^+$ and constructs a double spend tree $\widehat{\mathsf{DST}}$ with values (x_k, bh_k) where $k = 1, ..., n_k$. The exchange samples a uniformly random element r from $\mathbb{Z}_l^+$ and sets $\hat{C}_{res_k} = rG$. The kth simulated output is

$$\hat{z}_{n_k} = [\mathsf{bh}_k,\ \mathsf{KIT}_k.root,\ \mathsf{TXOT}_k.root,\ \widehat{\mathsf{DST}}_k.root, \hat{C}^{res}_{n_k}].$$

- The exchange sets $\mathcal{Z}_{sim} = \{\hat{z}_{n_k} \mid k = 1, \ldots, f(\lambda)\}$.

If there exists no PPT adversary $\mathcal{A}$ that can distinguish between $\mathcal{Z}_{\text{actual}}$ and $\mathcal{Z}_{sim}$ except with a negligible probability of success, then we say $\mathcal{Z}_{\text{actual}}$ does not reveal any information about the exchange. We define the privacy experiment `PORPriv` as follows.

1. The exchange sets $\mathcal{Z}_0 = \mathcal{Z}_{sim}$ and $\mathcal{Z}_1 = \mathcal{Z}_{\text{actual}}$.
2. The exchange chooses a bit b uniformly from $\{0, 1\}$.
3. The exchange sends $\mathcal{Z}_b$ to $\mathcal{A}$
4. $\mathcal{A}$ outputs a bit $b' = \mathcal{A}(\mathcal{Z}_b)$.
5. $\mathcal{A}$ succeeds if $b' = b$. Otherwise, it fails.

Definition B.2. A proof of reserves protocol is said to provide exchange privacy if for every PPT adversary $\mathcal{A}$ in the `PORPriv` experiment, there exists a negligible function $\mathsf{negl}(\lambda)$ of the security parameter λ such that,

$$Pr[b' = b] \leq \frac{1}{2} + \mathsf{negl}(\lambda).$$

THEOREM B.3. *The MProve-Nova proof of reserves protocol provides exchange privacy in the random oracle model.*

PROOF. We need to show that $\mathcal{A}$ cannot distinguish between $\mathcal{Z}_0$ and $\mathcal{Z}_1$ with a probability that is significantly greater than that of random guessing.

The outputs in $\mathcal{Z}_0$ and $\mathcal{Z}_1$ have the elements

$$(\mathsf{bh}_k,\ \mathsf{KIT}_k.root,\ \mathsf{TXOT}_k.root)_{k=1}^{f(\lambda)}$$

common between them. In addition to these elements, the simulated outputs in $\mathcal{Z}_0$ have

$$(\widehat{\mathsf{DST}}_k.root, \hat{C}^{res}_{n_k})_{k=1}^{f(\lambda)}$$

and the actual outputs in $\mathcal{Z}_1$ have

$$(\mathsf{DST}_k.root, C^{res}_{n_k})_{k=1}^{f(\lambda)}.$$

Consider the kth simulated commitment, $\hat{C}^{res}_{n_k} = rG$. Since r is uniformly random element from $\mathbb{Z}_l$, $\hat{C}^{res}_{n_k}$ is uniformly distributed over $\mathbb{G}$. Consider the kth actual commitment $C^{res}_{n_k}$. Since a uniformly random blinding factor is added to the actual reserve commitment at each step of the computation, $C^{res}_{n_k}$ is uniformly distributed over $\mathbb{G}$. Thus both $\hat{C}^{res}_{n_k}$ and $C^{res}_{n_k}$ are uniformly distributed over $\mathbb{G}$ for all $k \in \{1, 2, ..., f(\lambda)\}$.

Consider the kth simulated and actual DST roots $\widehat{\mathsf{DST}}_k.root$ and $\mathsf{DST}_k.root$, respectively. Both the roots are outputs of Poseidon hash H_{pos}. As H_{pos} is modeled as random oracle and assuming the private keys are generated uniformly at random, both $\widehat{\mathsf{DST}}_k.root$ and $\mathsf{DST}_k.root$ are uniformly distributed.

Now suppose an exchange generates two actual outputs $z_{n_l}, z_{n_m} \in \mathcal{Z}_1$, for the same blockheight bh and the same set of private keys. Then the DST roots in z_{n_l} and z_{n_m} will be same. Since two outputs in $\mathcal{Z}_1$ will have the same DST roots, adversary $\mathcal{A}$ can distinguish between $\mathcal{Z}_0$ and $\mathcal{Z}_1$. This case is eliminated with a restriction on the generation of the actual outputs. The exchange has to generate each actual output for a different blockheight. If the exchange chooses to generate the outputs at the same blockheight, then the set of private keys which the exchange uses to generate the output should not be the same. This restriction ensures that no two outputs in $\mathcal{Z}_1$ have the same DST root.

The above arguments show that $\mathcal{Z}_0$ and $\mathcal{Z}_1$ are computationally indistinguishable from each other and a PPT adversary $\mathcal{A}$ cannot distinguish between them with a probability that is significantly greater than that of random guessing. □

B.2.2 Proof of Non-collusion. Suppose two exchanges Ex1 and Ex2 have generated the proof of reserves for the same blockheight bh and Ex1 wants to prove non-collusion with Ex2. Ex2 has already made the values $v_j = H_{pos}(x_j^{(2)}||\mathsf{bh})$ which were inserted in its $\mathsf{DST}^{\mathsf{Ex2}}$ publicly available.

As discussed before, the non-collusion protocol does use publicly available data and any user can check non-collusion. But we proposed a protocol based on Nova for faster verification of non-collusion. Thus the protocol has no effect on the privacy of the exchange proving non-collusion i.e. Ex1. But Ex2 has to reveal the values $\left(v_j\right)_{j=1}^{n}$ to the public. In this section, we will argue that this has no effect on the privacy of Ex2.

We consider the scenario when an exchange generates $f(\lambda)$ number of proofs of reserves, where $f(\lambda)$ denotes a polynomial of the security parameter λ. Suppose for each proof of reserves, the exchange inserted a polynomial $g(\lambda)$ number of values in the DST. Thus all the values released publicly by the exchange for proof of non-collusion will be upper bounded by a polynomial $h(\lambda) = f(\lambda) * g(\lambda)$. Let us collect all these values in a vector $V_{\text{actual}} = (v_j)_{j=1}^{h(\lambda)}$.

We want to show that V_{actual} does not reveal any information about the exchange. All the values in V_{actual} are Poseidon hashes and map to a finite field $\mathbb{F}_q$, i.e. $v_j \in \mathbb{F}_q$ for all $j \in \{1, 2, ..., h(\lambda)\}$.

The exchange generates a vector with simulated values $V_{sim} = (\hat{v}_j)_{j=1}^{h(\lambda)}$ such that $\hat{v}_j \overset{\$}{\leftarrow} \mathbb{F}_q$, where $\overset{\$}{\leftarrow}$ denotes uniform random sampling. If there exists no PPT adversary $\mathcal{A}$ that can distinguish between V_{actual} and V_{sim} except with a negligible probability of success, then we say V_{actual} does not reveal any information about the exchange. Now similar to PORPriv experiment we define PNCPriv experiment as follows.

(1) The exchange sets $V_0 = V_{sim}$ and $V_1 = V_{\text{actual}}$.
(2) The exchange chooses a bit b uniformly from $\{0, 1\}$.
(3) The exchange sends V_b to $\mathcal{A}$
(4) $\mathcal{A}$ outputs a bit $b' = \mathcal{A}(V_b)$.
(5) $\mathcal{A}$ succeeds if $b' = b$. Otherwise, it fails.

Definition B.4. V_{actual} does not reveal any information about the exchange if for every PPT adversary $\mathcal{A}$ in the PNCPriv experiment, there exists a negligible function $\mathsf{negl}(\lambda)$ of the security parameter λ such that,

$$Pr[b' = b] \leq \frac{1}{2} + \mathsf{negl}(\lambda).$$

Theorem B.5. *V_{actual} reveals no information about the exchange in the random oracle model.*

Proof. We need to show that $\mathcal{A}$ cannot distinguish between V_0 and V_1 with a probability that us non-negligibly greater than that of random guessing.

We know that all the $f(\lambda)$ proof of reserves correspond to a different blockheight of the Monero blockchain. If the exchange uses the same blockheight for different outputs then the set of private keys used to generate proof of reserves should not be same. Thus either x_j or blockheight bh corresponding to each $v_j \in V_1$ is different. Thus no two values in V_1 are identical.

As H_{pos} is modeled as random oracle, v_j is uniformly distributed. Thus each v_j is indistinguishable from a uniformly random elements of $\mathbb{F}_q$, i.e. elements of V_0. Also since no two values in V_1 are same, V_1 is computationally indistinguishable from V_0. Thus a PPT adversary $\mathcal{A}$ cannot distinguish between them with a probability that is non-negligibly greater than that of random guessing. □

B.3 Collusion Resistance

The collusion resistance property prevents two exchanges from colluding to generate the proof of reserves. This property follows from the construction of our non-collusion protocol. The protocol guarantees to detect collusion provided that the two exchanges giving the proof of non-collusion have generated the proof of reserves at the same blockheight.