

Hash Function Attacks

Technical Field

- 5 The present disclosure relates to a form of zero-knowledge proof which may be used to prove the existence of a successful attack on a hash function, without revealing details of the attack itself.

Background

10

A blockchain refers to a form of distributed data structure, wherein a duplicate copy of the blockchain is maintained at each of a plurality of nodes in a peer-to-peer (P2P) network. The blockchain comprises a chain of blocks of data, wherein each block comprises one or more transactions. Each transaction may point back to a preceding transaction in a sequence.

- 15 Transactions can be submitted to the network to be included in new blocks. New blocks are created by a process known as “mining”, which involves each of a plurality of mining nodes competing to perform “proof-of-work”, i.e. solving a cryptographic puzzle based on a pool of the pending transactions waiting to be included in blocks.

- 20 Conventionally the transactions in the blockchain are used to convey a digital asset, i.e. data acting as a store of value. However, a blockchain can also be exploited in order to layer additional functionality on top of the blockchain. For instance, blockchain protocols may allow for storage of additional user data in an output of a transaction. Modern blockchains are increasing the maximum data capacity that can be stored within a single transaction,
25 enabling more complex data to be incorporated. For instance this may be used to store an electronic document in the blockchain, or even audio or video data.

- Each node in the network can have any one, two or all of three roles: forwarding, mining and storage. Forwarding nodes propagate transactions throughout the nodes of the
30 network. Mining nodes perform the mining of transactions into blocks. Storage nodes each store their own copy of the mined blocks of the blockchain. In order to have a transaction

recorded in the blockchain, a party sends the transaction to one of the nodes of the network to be propagated. Mining nodes which receive the transaction may race to mine the transaction into a new block. Each node is configured to respect the same node protocol, which will include one or more conditions for a transaction to be valid. Invalid transactions
5 will not be propagated nor mined into blocks. Assuming the transaction is validated and thereby accepted onto the blockchain, then the transaction (including any user data) will thus remain stored at each of the nodes in the P2P network as an immutable public record.

The miner who successfully solved the proof-of-work puzzle to create the latest block is
10 typically rewarded with a new transaction called a “generation transaction” which generates a new amount of the digital asset. The proof-of work incentivises miners not to cheat the system by including double-spending transactions in their blocks, since it requires a large amount of compute resource to mine a block, and a block that includes an attempt to double spend is likely not be accepted by other nodes.

15 In an “output-based” model (sometimes referred to as a UTXO-based model), the data structure of a given transaction comprises one or more inputs and one or more outputs. Any spendable output comprises an element specifying an amount of the digital asset, sometimes referred to as a UTXO (“unspent transaction output”). The output may further
20 comprise a locking script specifying a condition for redeeming the output. Each input comprises a pointer to such an output in a preceding transaction, and may further comprise an unlocking script for unlocking the locking script of the pointed-to output. So consider a pair of transactions, call them a first and a second transaction. The first transaction comprises at least one output specifying an amount of the digital asset, and comprising a
25 locking script defining one or more conditions of unlocking the output. The second transaction comprises at least one input, comprising a pointer to the output of the first transaction, and an unlocking script for unlocking the output of the first transaction.

In such a model, when the second transaction is sent to the P2P network to be propagated
30 and recorded in the blockchain, one of the conditions for validity applied at each node will be that the unlocking script meets all of the one or more conditions defined in the locking

script of the first transaction. Another will be that the output of the first transaction has not already been redeemed by another, earlier valid transaction. Any node that finds the second transaction invalid according to any of these conditions will not propagate it nor include it for mining into a block to be recorded in the blockchain.

5

An alternative type of transaction model is an account-based model. In this case each transaction does not define the amount to be transferred by referring back to the UTXO of a preceding transaction in a sequence of past transactions, but rather by reference to an absolute account balance. The current state of all accounts is stored by the miners separate to the blockchain and is updated constantly. A transaction of an account-based model can also include smart contract which runs at each node at the same time as validating the transaction.

10

A transaction may include a digital signature to as evidence an identity attached to the transaction. As will be familiar to a person skilled in the art, a cryptographic signature can be generated based on a private key V and verified based on a corresponding public key P of a private-public key pair. Given a signature generated by applying the private key V to a message m , it is possible for another party to verify using P that the signature was generated using V , without that party knowing V (hence verifying a signature itself is another form of knowledge proof in its own right).

15

20

One form of algorithm for this is an elliptic curve digital signature algorithm (ECDSA) which operates based on elliptic curve cryptography (ECC). In this case P and V are related by:

25

$$P = V \cdot G$$

where P is a two-element vector (P_x, P_y) , V is a scalar, and G is a two element vector (G_x, G_y) representing a predetermined point on a two-dimensional elliptic curve (the "generator point"). The operation " $\cdot$ " is the scalar elliptic curve multiplication - a known form of operation that translates from one point on an elliptic curve to another.

30

The ECDSA signature is a pair of elements (r, s) commonly known in the art as the r-part (r) and the s-part (s) respectively. The signature (r, s) is generated by applying the private key V to a message m . In the case of a transaction for recordal on a blockchain, m will be a part of the transaction and the signature will be tagged onto the transaction in addition to that part in the clear, to enable that transaction to be validated. For instance in an output-based model, the signature signs a part of Tx_2 and is included in the locking script of Tx_2 in order to unlock the output of Tx_1 . The signed part typically includes the output(s) of the transaction so these cannot be changed without invalidating the signature and therefore the transaction.

Whatever transaction model is used, the signature (r, s) is computed as:

$$r = [R]_x, \text{ where } R = k \cdot G$$

$$s = k^{-1}(H_{sig}(m) + rV) \bmod n$$

where $[R]_x$ denotes the x-coordinate of the two element vector $R = (R_x, R_y)$. k is known as the ephemeral key. It is selected from a set $[1, n - 1]$, typically at random, where n is a prime modulus and $[1, n - 1]$ is the set of real integers in the range 1 to $n - 1$ inclusive. H_{sig} is a hash function.

Given knowledge of the signature (r, s) , the message m , and the public key P , it is possible for any party who does not know the private key V to verify that the signature was generated using the private key to V to the message m . This is done by computing:

$$R' = H_{sig}(m)s^{-1} \cdot G + rs^{-1} \cdot P$$

and verifying that $[R']_x = r$. The signature is verified only if this is true, but is not verified otherwise. This can be taken as a verification that the party associated with the public key P was indeed the signer of the message.

Cryptographic hash functions (such as H in the above) form the basis of many critical cryptography functions. For example, they are applied extensively in numerous encryption and digital signature systems in use today. For example, in a blockchain context, such hash functions form part of the processes used to generate and verify digital signatures applied to transactions, e.g. as part of the ECDSA as outlined above. SHA-256 is an example of a cryptographic hash function widely used in blockchain systems. A cryptographic hash function is a one-way function which maps a piece of data of arbitrary size to a fixed-length bit string. The bit string returned by the hash function is referred to a “hash value” (or, equivalently, an “image”) and any piece of data for which the hash function returns a given hash value is called a preimage of that hash value (with respect to the hash function in question).

One-way implies the function should be easy to compute but hard to reverse, where the terms easy and hard are used in the specific cryptography sense. That is, given a preimage (d), it should be computationally feasible to compute the cryptographic hash function (H) on that preimage and thereby obtain a corresponding hash value ($h = H(d)$); in practice, this generally implies that computations can be completed in polynomial time. However, given a hash value (h) and a hash function (H), it should be computationally infeasible to recover any preimage to that hash function (i.e. any d such that $H(d) = h$), e.g. because the reverse computations could only be completed in exponential time). Polynomial time and exponential time denote the manner in which the number of operations required to complete an algorithm scales in relation to the size of the input to that algorithm – for a sufficiently large input, exponential time algorithms are practically infeasible to compute.

25

Summary

In a cryptography context, an “preimage attack” on a cryptographic hash function refers to an attempt to recover at least one preimage to a specified hash value. The hash value could be specified directly, or by reference to a given piece of data in which case the aim is to find a second piece of data on which the hash function returns the same hash value as the first

30

piece of data (a so-called “second preimage attack”). A “collision attack” refers to an attempt to recover at least two preimages of the same hash value – i.e. two pieces of data (d_1, d_2) which are different from each other, but for which the hash function returns the same hash value (i.e. $H(d_1) = H(d_2) = h$). For a collision attack, the hash value h may or may not be specified (note that, under this definition, a second preimage attack is a form of collision attack in which the hash value is specified in advance). A successful preimage or collision attack on a cryptographic hash function may denote a vulnerability in any cryptographic system which is secured based on that hash function. For example, in a blockchain context, such attacks, if successful, could be used as a basis for digital signature forgery, allowing an attacker to create transactions with valid signatures without knowledge of the corresponding public key. Such attacks can also be used as a basis for forgery in other contexts, such as digital signature forgery.

The implications of a successful attack on cryptographic hash function are concrete and potentially severe. For example, in December 2008, a group of security researchers successfully forged a X.509 public key certificate, exploiting a collision attack discovered in respect of the MD5 hash function (see “MD5 considered harmful today - Creating a rogue CA certificate” Alexander Sotirov et. al; December 30, 2008). MD5 was originally designed to be used as cryptographic hash function, but its security has now been severely compromised. As a consequence, MD5 has been widely discredited in the field of cryptography and the use of MD5 for cryptographic purposes has all but discontinued (it is still used for non-cryptographic purposes e.g. in distributed databases). As well as serving to illustrate the practical consequences of a successful hash function attack, the history of MD5 also serves to illustrate a point of terminology adopted in the present disclosure: unless otherwise indicated, the term “cryptographic hash function” is used herein as a shorthand to mean an *assumed* cryptographic hash function, i.e. which is *assumed* to have good cryptographic properties; however, a successful attack on the hash function may expose that assumption as false.

An “attacker” in the present context need not have malicious intent. Benevolent attacks may be attempted on a hash function to try to discover security vulnerabilities, and take

steps to mitigate them, before they are discovered by a malicious attacker. In that context, there are significant benefits in being able to communicate successful attacks on a hash function in robust and verifiable manner. The present disclosure provides a mechanism by which a successful preimage or collision attack on a hash function can be proved by way of a blockchain transaction, without revealing the preimage or preimages discovered in the attack (i.e. the proof is “zero-knowledge” in the cryptography sense). The present disclosure recognizes that a blockchain network is a highly-suitable decentralized platform for safely distributing knowledge between cryptography experts and other interested parties and recording that knowledge immutably.

A prover constructs a zero-knowledge proof (ZKP) on the *r-part* (r) of an ECDSA signature (r, s) to be included in a proof transaction. The ZKP is embodied as a ZKP component (π) of the proof transaction in conjunction with the *r-part* (r) of the ECDSA signature (r, s) itself.

A blockchain network node which receives the proof transaction for processing:

- (i) verifies the signature (r, s) of the proof transaction based on an associated public key P and a signed part m of the proof transaction, and
- (ii) applies a proof verification algorithm ($\mathcal{V}$) to determine whether the ZKP component π of the proof transaction is correct for the *r-part* (r) of the verified signature (r, s) and a defined hash value h and a defined hash function H , in the sense of proving the following statement:

“an ephemeral key counterpart to the *r-part* (r) of the signature (r, s) is a preimage of the defined hash value h with respect to the defined hash function H ”.

In other words, proof that the prover knows a secret ephemeral key k such that

$H(k) = h$ and $R = k \cdot G$ where $r = [R]_x$. The proof is zero-knowledge in that the pre-image d itself is not revealed in the proof transaction.

Thus, as well as providing the usual benefits of an ECDSA signature, the signature itself also acts in conjunction with the ZKP component to prove a successful attack on the hash function H . This has various advantages which are set out in further detail below with reference to the described embodiments. One particular advantage of constructing the proof on the *r-part* of the signature is that it avoids any requirement to have any pre-

specified identity attached to the proof transaction, i.e. there is no requirement for the transaction to be signed using a pre-specified private key or, equivalently, no requirement for the transaction to include a signature associated with a pre-specified public key; the signature can be generated using any private key V provided the signature is valid and the
5 ZKP component π is correct for the r-part of that signature in the above sense.

In embodiments of the present disclosure, the proof transaction provides a zero-knowledge proof of a successful hash *collision* attack on the defined hash function H . To this end, the proof transaction may comprise first and second ECDSA signatures (r_1, s_1) , (r_2, s_2) together
10 with first and second ZKP components π_1, π_2 . In this case, a proof verification algorithm $\mathcal{V}$ may be constructed to verify that the r-parts r_1, r_2 of the first and second signatures have been generated using *different* ephemeral keys k_1, k_2 but for which the defined hash function H returns the *same* hash value h (the two requirements for a collision; k_1 and k_2 are said of be colliding preimages in that event). A particular benefit in respect of hash
15 collision attacks is that, as well as keeping the colliding preimages k_1, k_2 secret, *the hash value h need not be known in advance* (and thus need not be specified in a challenge transaction, any other transaction, anywhere else in the blockchain, or otherwise).

The fact that the pre-image(s) need not be revealed is important for security: if the
20 preimage(s) were to be revealed in the transaction, another attacker who is malicious and has access to the proof transaction could immediately replicate the attack in a real-world context (e.g. using it as a basis for digital signature forgery). Moreover, in the case that the challenge transaction provides a digital asset in return for a correct proof, a malicious party (e.g. miner) could claim that digital asset for himself now that he knows the preimage(s).
25 However, a zero-knowledge proof prevents this, whilst still providing valuable information about any vulnerabilities in a system that is secured based on the hash function and that allows action to be taken to mitigate those vulnerabilities; proof of the existence of the attack, even without knowledge of the preimage(s), is still highly valuable in a security context as it can reveal the existence of security vulnerabilities which can now be mitigated,
30 whereas a malicious attacker with access to the transaction simply knows that the attack exists but does not know how to replicate it.

A first aspect of the present disclosure provides a computer-implemented method of processing at least one proof transaction at a node of a blockchain network, the method comprising:

- receiving the at least one proof transaction, wherein the at least one proof transaction comprises at least one Elliptic Curve Digital Signature Algorithm (ECDSA) signature and at least one zero-knowledge proof (ZKP) component;
- verifying the ECDSA signature of the at least one proof transaction based on a public key associated with the ECDSA signature and a signed part of the at least one proof transaction; and
- determining whether at least one ZKP component proves knowledge of a successful attack on the hash function by determining whether the ZKP component is correct for the ECDSA signature and a defined hash value and a defined hash function, in that it proves an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of the defined hash value with respect to the defined hash function

Brief Description of the Drawings

To assist understanding of embodiments of the present disclosure and to show how such embodiments may be put into effect, reference is made, by way of example only, to the accompanying drawings in which:

Figure 1 is a schematic block diagram of a system for implementing a blockchain,

Figure 2 schematically illustrates some examples of transactions which may be recorded in a blockchain,

Figure 3 is a schematic block diagram of another system for implementing a blockchain,

Figure 4 is a schematic block diagram of a piece of node software for processing transaction in accordance with a node protocol of an output-based model,

Figure 5 is a schematically illustrates an example set of transactions,

Figures 6A-6D schematically illustrate some of the principles behind an elliptic curve digital signature algorithm (ECDSA),

5

Figure 7 is a schematic illustration of one possible implementation of a type of knowledge proof referred to herein as an r-puzzle (or synonymously an r-challenge),

Figure 8 is a schematic illustration of another possible implementation of an r-puzzle,

10

Figure 9 is a schematic illustration of another possible implementation of an r-puzzle,

Figure 10 is a schematic illustration of yet another possible implementation of an r-puzzle,

15 Figure 11 is a schematic block diagram of a piece of node software for processing transactions in accordance with a node protocol of an account-based model,

Figure 12 schematically illustrates an example format for an ECDSA signature,

20 Figure 13 is a step-by-step script analysis of an example implementation of a locking and unlocking script for one form of r-puzzle,

Figure 14 shows a hash collision bounty transaction and a corresponding proof transaction,

25 Figure 15 shows (a) an arithmetic circuit gate and (b) an arithmetic circuit,

Figure 16 shows another example arithmetic circuit, and

Figure 17 shows a process flow for a method of verifying an arithmetic circuit proof.

30

Detailed Description of Embodiments

In some cryptographic schemes a verifier may require convincing that a person (called a prover or challengee) has some piece of information in what is called a *knowledge proof*. This may be done by providing the piece of information directly to the verifier, or the prover may be required to perform a calculation requiring the piece of information in what is sometimes called a *zero-knowledge proof* (so called because in this type of knowledge proof, the verifier themselves does not need to know the piece of information to set the challenge and nor does the piece of information need to be revealed to the verifier in order to verify that the prover knows the piece of information). For a computational method, a verification calculation must be performed on the input data. A straightforward way of proving knowledge of a secret value is through use of a cryptographic hash function due to its features of preimage and collision resistance. This hash method can be easily integrated in many blockchain applications as hash functions form a fundamental part of their private key-public key cryptosystem. This type of knowledge proof is so prolific in blockchain applications that is typically referred to as a *hash puzzle*.

In UTXO-based blockchains, the solution to a “hash puzzle” (the preimage of the hashed value) can be set as a spending condition so the verification is performed by the miner as part of the transaction verification. If an input of a second transaction Tx_2 points to this output of Tx_1 , then the unlocking script in that input of Tx_2 will have to solve the hash puzzle in order to successfully redeem the output of Tx_1 . The hash puzzle comprises a hash value h , being a hash of d , i.e. $h = H_{puz}(d)$. The puzzle also comprises a piece of script that, when run at a node together with the unlocking script of Tx_2 , will take a data value d' purporting to be d from the unlocking script of Tx_2 , hash it with the hash function H_{puz} , and compare with the hash value h included in the locking script of Tx_1 . I.e. it checks whether $h = H_{puz}(d')$ and will only unlock the output of Tx_1 if the result of the comparison is yes (or “true” in the terminology of the art). Thus the beneficiary of Tx_2 can only unlock the output of Tx_1 if d is included in the unlocking script of Tx_2 to prove knowledge of d .

An issue with using the conventional hash puzzle alone is that an unscrupulous miner or other node could observe d in the unlocking script of Tx_2 , then create and mine (or publish)

his own version Tx_2^* of Tx_2 , paying himself in the output of Tx_2^* instead of the intended recipient (e.g. Bob) as in Tx_2 . The existing way to avoid this is to additionally include a “pay-to-public key hash” (P2PKH) requirement in the locking script of Tx_1 . In addition to the knowledge proof for d , this requires a cryptographic signature of the intended payee to be
5 included in the unlocking script of Tx_2 .

A hash puzzle and P2PKH can also be implemented using a smart contract in an account-based model, rather than the locking and unlocking scripts of an output-based model.

10 However, an issue with using a hash puzzle as the knowledge proof, combined with P2PKH to avoid the possible miner attack, is that while the P2PKH ensures the payment is only made to the party who knew d in the first place, it also means that the output of Tx_1 is tied to a specific predetermined recipient or set of recipients (it is possible to specify alternative conditions which could include alternative recipients, but they still have to be pre-
15 identified).

In the present disclosure, a knowledge proof is provided that circumvents these issues while still allowing for the validation to be performed by the miner (or a forwarding node). To do this, the knowledge proof is connected to an ephemeral key corresponding to an elliptic
20 curve digital signature algorithm (ECDSA) signature. As the cryptographic primitives used in this algorithm are native to many blockchains, it can be readily integrated into the current infrastructure.

ATTACKS ON HASH FUNCTIONS - OVERVIEW

25 The described embodiments use blockchain technology as a basis for communicating and recording attacks on hash functions. Proof of a successful attack on a hash function is submitted to a node of a blockchain network for processing in a transaction. The node processes the transaction to determine whether the submitted transaction does indeed
30 prove that a successful attack has taken place. This processing may be performed implicitly as part of a transaction validation process applied to the transaction. A transaction which

does provide such proof is stored in a blockchain maintained by the blockchain network (assuming it meets any validity requirements imposed by the blockchain network in question) and thus provides an immutable record evidencing the successful attack. This, in turn, can be used by anyone with access to the blockchain e.g. as a basis for mitigating action, with the aim of mitigating any security vulnerabilities exposed by the successful attack.

As indicated above, the proof is a zero-knowledge proof (ZKP). A ZKP means a proof that a party knows a value (in this context, the preimage or preimages) without revealing the value itself (i.e. in this context, without revealing the preimage or preimages).

It is also desirable in a blockchain context to be able to attribute an identity to a party which has discovered a successful hash function attack, and this can be achieved by providing a valid digital signature in the transaction which signs a part of the transaction – in that event, an identity is attached to the transaction as “the party who holds a private key counterpart to a public key associated with the digital signature”.

However, the mere inclusion of a digital signature cannot *per se* protect against a situation in which a malicious miner extracts the proof of the hash attack from the transaction, and includes it in a new transaction signed with its own private key and which thus has a different identity attached. What can, however, mitigate this situation is providing an element of the proof in the signature itself, such that the proof is no longer correct (valid) if the signature is changed.

The described embodiments leverage the fact that it is possible to construct a zero-knowledge proof of the following form:

“the preimage of hash value h with respect to hash function H is equal to a scalar multiplier d for which $P = d \cdot G$ ”

This can be equivalently formulated as a proof that $H(d) = h$ and $d \cdot G = P$. Here, “ $\cdot$ ” denotes elliptic curve scalar multiplication on a defined elliptic curve $\mathcal{E}_n$ with a defined generator point G . The proof is zero-knowledge in that it can robustly prove the above

statement *without revealing* d , i.e. the preimage of the hash function is secret in the cryptographic sense; in order to prove knowledge of d , a party can use d to derive the corresponding point P on the elliptic curve, and present P along with a zero-knowledge proof component π to prove their knowledge of d (instead of providing d itself). The only
 5 public information needed to verify (i.e. determine the correctness of) the ZKP component proof π is $\mathcal{E}_n, G, P, h$ and H .

Note: the hash function H for which the proof is constructed may be the same or different than the hash function H_{sig} used to generate and verify the ECDSA signature. Although the
 10 notation H may at times be used in the following description to represent the hash function used to generate the signature (i.e. H_{sig}), this is purely for the sake of simplicity and does not necessarily imply any equality with the hash function for which the proof is constructed.

One way of constructing the zero-knowledge proof on the above would be to use the
 15 preimage d to derive the *public* key, i.e. $P = d \cdot G$ used to generate an ECDSA signature. This can be equivalently stated as signing the message m using the preimage d as the *private* key. Consider a transaction $Tx = (m, P, (r, s), \pi)$, i.e. a transaction comprising a signed part m (the message), an ECDSA signature (r, s) associated with a public key P and a zero-knowledge proof component π . With this construction, the zero-knowledge proof
 20 component π would be verified based on P , i.e. to determine whether the zero-knowledge proof component π does indeed prove the private key counterpart to P to be a preimage of the hash function. A consequence of this is that a malicious miner cannot simply include π in a new transaction with a new signature: the new signature will be associated with a different public key P' , and the zero-knowledge proof component π will not be correct for
 25 P' (because the corresponding private key is no longer a preimage of h).

However, the present disclosure recognizes a problem with such an approach in that it restricts the identity of the party who can provide the transaction to the holder of the private key d .

The present disclosure further recognizes that an ECDSA signature in fact relies on *two* instances of elliptic curve scalar multiplication, and *two* secret values, to generate a signature for a given message m : one to derive the public key as $P = V \cdot G$ (from the private key V) and the other to derive the public ephemeral key as $R = k \cdot G$ (from the secret ephemeral key k). Rather than construct the proof on the public key P , the present disclosure instead constructs the zero-knowledge proof on the *r-part* (r) of the signature (r, s) , i.e. the *r-part* is derived as:

$$R = d \cdot G$$

$$r = [R]_x$$

$$H(d) = h$$

With this new construction, given a transaction $Tx = (m, P, (r, s), \pi)$, the ZKP component π is correct if it proves the following statement:

“the preimage of hash value h with respect to hash function H is equal to the scalar multiplier d used to derive the *r-part* r of the signature (r, s) ”

which in turn can be verified based on $\mathcal{E}_n, G, h, H$ and r . This can be equivalently phrased as proof that the ephemeral key k used to derive the *r-part* r is a preimage of h , i.e. $k = d$ where $H(d) = h$. As a consequence, there is now no restriction on the private key V used to generate the signature above and beyond the requirement for the signature to be valid and the proof on the *r-part* to be correct.

The term “*r-puzzle*” is used herein in relation to any knowledge proof constructed on the *r-part* of an ECDSA signature. An *r-puzzle* may be embodied in code of a challenge transaction as one or more proof requirement which must be met by the *r-part* of at least one ECDSA signature in a proof transaction (if all of the requirement(s) are met, the proof transaction is said to “solve” the *r-puzzle* of the challenge transaction). The requirement(s) on the *r-part* impose implicit requirement(s) on the ephemeral key k used to derive the *r-part*, hence a solution to the *r-puzzle* proves the a challengee (prover) to be in possession of an ephemeral key k which satisfies those requirement(s), without revealing the ephemeral key k itself. Under this definition, the requirement to provide a correct proof of the statement set out in the preceding paragraph is a particular form of *r-puzzle* by virtue of the fact that the zero-knowledge proof is constructed on the *r-part* of the signature.

In the specific context of proving knowledge of a successful attack on a cryptographic hash function, the r-puzzle requires a proof transaction to contain (i) at least one valid signature (r, s) and (ii) at least one zero-knowledge proof component π which is correct for the r-part r of that signature and the defined hash h value and hash function H .

5

In the more specific context of proving knowledge of a successful hash function attack, the r-puzzle requires at least two valid signatures (r_1, s_1) and (r_2, s_2) and at least two zero-knowledge proof components π_1, π_2 where π_1 and π_2 are correct for r_1 and r_2 respectively and for the same hash value h and the same hash function H as each other.

10

As noted above, although the teaching of the present disclosure applies more generally, there are nevertheless particular benefits in respect of hash collision attacks, in the fact that neither the hash value h nor the colliding pre-images k_1 and k_2 (i.e. k_1 and k_2 where $k_1 \neq k_2$ and $H(k_1) = H(k_2) = h$) need to be known in advance of setting the r-puzzle.

15

In order to incentive benevolent attackers to provide proof of a successful attack on a hash function, a “bounty” may be offered in return for a transaction which solves the corresponding r-puzzle (the proof transaction). In a UTXO model, the bounty can be embodied in a locking script of one or more unspent transaction outputs, which in turn can only be “redeemed” (spent) by a transaction containing unlocking script which solves the r-puzzle. In this context, a benefit of the above proof construction is that a malicious miner cannot alter the transaction to redeem the unspent transaction outputs (the bounty) for himself.

25 In an account-based model, a smart contract can be coded to release an amount of digital asset from an account of the smart contract as a bounty, in return for a transaction which solves the r-puzzle.

A transaction which can be redeemed in return for such a transaction may be referred to
30 herein as a “bounty” transaction – which is an example of a “challenge” transaction as that term is used herein.

A “hash collision bounty” transaction requires provide of a hash collision attack on a defined hash function. Such a transaction need not specify the hash value h (nor is it necessary to specify the hash value h in any other manner in advance) - in which case, all that is required is proof that a party knows two preimages which return the same hash value as each other.

A “hash preimage bounty” transaction requires proof of preimage attack on a defined hash function, which may be defined in the hash preimage bounty transaction (e.g. in its locking script) or by a related transaction (e.g. in an account-based model, h could be uploaded to a smart contract register in an intermediate transaction other than the transaction which creates the smart contract account).

Further details are described below, but first some additional context to some of the example embodiments is provided.

15

EXAMPLE SYSTEM OVERVIEW

Figure 1 shows an example system 100 for implementing a blockchain 150. The system 100 comprises a packet-switched network 101, typically a wide-area internetwork such as the Internet. The packet-switched network 101 comprises a plurality of nodes 104 arranged to form a peer-to-peer (P2P) overlay network 106 within the packet-switched network 101. Each node 104 comprises computer equipment of a peers, with different ones of the nodes 104 belonging to different peers. Each node 104 comprises processing apparatus comprising one or more processors, e.g. one or more central processing units (CPUs), accelerator processors, application specific processors and/or field programmable gate arrays (FPGAs). Each node also comprises memory, i.e. computer-readable storage in the form of a non-transitory computer-readable medium or media. The memory may comprise one or more memory units employing one or more memory media, e.g. a magnetic medium such as a hard disk; an electronic medium such as a solid-state drive (SSD), flash memory or EEPROM; and/or an optical medium such as an optical disk drive.

The blockchain 150 comprises a chain of blocks of data 151, wherein a respective copy of the blockchain 150 is maintained at each of a plurality of nodes in the P2P network 160. Each block 151 in the chain comprises one or more transactions 152, wherein a transaction in this context refers to a kind of data structure. The nature of the data structure will
5 depend on the type of transaction protocol used as part of a transaction model or scheme. A given blockchain will typically use one particular transaction protocol throughout. In one common type of transaction protocol, the data structure of each transaction 152 comprises at least one input and at least one output. Each output specifies an amount representing a quantity of a digital asset belonging to a user 103 to whom the output is cryptographically
10 locked (requiring a signature of that user in order to be unlocked and thereby redeemed or spent). Each input points back to the output of a preceding transaction 152, thereby linking the transactions.

At least some of the nodes 104 take on the role of forwarding nodes 104F which forward
15 and thereby propagate transactions 152. At least some of the nodes 104 take on the role of miners 104M which mine blocks 151. At least some of the nodes 104 take on the role of storage nodes 104S (sometimes also called “full-copy” nodes), each of which stores a respective copy of the same blockchain 150 in their respective memory. Each miner node 104M also maintains a pool 154 of transactions 152 waiting to be mined into blocks 151. A
20 given node 104 may be a forwarding node 104, miner 104M, storage node 104S or any combination of two or all of these.

In a given present transaction 152j, the (or each) input comprises a pointer referencing the output of a preceding transaction 152i in the sequence of transactions, specifying that this
25 output is to be redeemed or “spent” in the present transaction 152j. In general, the preceding transaction could be any transaction in the pool 154 or any block 151. The preceding transaction 152i need not necessarily exist at the time the present transaction 152j is created or even sent to the network 106, though the preceding transaction 152i will need to exist and be validated in order for the present transaction to be valid. Hence
30 “preceding” herein refers to a predecessor in a logical sequence linked by pointers, not necessarily the time of creation or sending in a temporal sequence, and hence it does not

necessarily exclude that the transactions 152i, 152j be created or sent out-of-order (see discussion below on orphan transactions). The preceding transaction 152i could equally be called the antecedent or predecessor transaction.

5 The input of the present transaction 152j also comprises the signature of the user 103a to whom the output of the preceding transaction 152i is locked. In turn, the output of the present transaction 152j can be cryptographically locked to a new user 103b. The present transaction 152j can thus transfer the amount defined in the input of the preceding transaction 152i to the new user 103b as defined in the output of the present transaction
10 152j. In some cases a transaction 152 may have multiple outputs to split the input amount between multiple users (one of whom could be the original user 103a in order to give change). In some cases transaction can also have multiple inputs to gather together the amounts from multiple outputs of one or more preceding transactions, and redistribute to one or more outputs of the current transaction.

15

The above may be referred to as an “output-based” transaction protocol, sometimes also referred to as an unspent transaction output (UTXO) type protocol (where the outputs are referred to as UTXOs). A user’s total balance is not defined in any one number stored in the blockchain, and instead the user needs a special “wallet” application 105 to collate the
20 values of all the UTXOs of that user which are scattered throughout many different transactions 152 in the blockchain 151.

An alternative type of transaction protocol may be referred to as an “account-based” protocol, as part of an account-based transaction model. In the account-based case, each
25 transaction does not define the amount to be transferred by referring back to the UTXO of a preceding transaction in a sequence of past transactions, but rather by reference to an absolute account balance. The current state of all accounts is stored by the miners separate to the blockchain and is updated constantly. In such a system, transactions are ordered using a running transaction tally of the account (also called the “position”). This value is
30 signed by the sender as part of their cryptographic signature and is hashed as part of the transaction reference calculation. In addition, an optional data field may also be signed the

transaction. This data field may point back to a previous transaction, for example if the previous transaction ID is included in the data field.

With either type of transaction protocol, when a user 103 wishes to enact a new transaction 152j, then he/she sends the new transaction from his/her computer terminal 102 to one of the nodes 104 of the P2P network 106 (which nowadays are typically servers or data centres, but could in principle be other user terminals). This node 104 checks whether the transaction is valid according to a node protocol which is applied at each of the nodes 104. The details of the node protocol will correspond to the type of transaction protocol being used in the blockchain 150 in question, together forming the overall transaction model. The node protocol typically requires the node 104 to check that the cryptographic signature in the new transaction 152j matches the expected signature, which depends on the previous transaction 152i in an ordered sequence of transactions 152. In an output-based case, this may comprise checking that the cryptographic signature of the user included in the input of the new transaction 152j matches a condition defined in the output of the preceding transaction 152i which the new transaction spends, wherein this condition typically comprises at least checking that the cryptographic signature in the input of the new transaction 152j unlocks the output of the previous transaction 152i to which the input of the new transaction points. In some transaction protocols the condition may be at least partially defined by a custom script included in the input and/or output. Alternatively it could simply be fixed by the node protocol alone, or it could be due to a combination of these. Either way, if the new transaction 152j is valid, the current node forwards it to one or more others of the nodes 104 in the P2P network 106. At least some of these nodes 104 also act as forwarding nodes 104F, applying the same test according to the same node protocol, and so forward the new transaction 152j on to one or more further nodes 104, and so forth. In this way the new transaction is propagated throughout the network of nodes 104.

In an output-based model, the definition of whether a given output (e.g. UTXO) is spent is whether it has yet been validly redeemed by the input of another, onward transaction 152j according to the node protocol. Another condition for a transaction to be valid is that the

output of the preceding transaction 152i which it attempts to spend or redeem has not already been spent/redeemed by another valid transaction. Again if not valid, the transaction 152j will not be propagated or recorded in the blockchain. This guards against double-spending whereby the spender tries to spend the output of the same transaction more than once. An account-based model on the other hand guards against double-spending by maintaining an account balance. Because again there is a defined order of transactions, the account balance has a single defined state at any one time.

In addition to validation, at least some of the nodes 104M also race to be the first to create blocks of transactions in a process known as mining, which is underpinned by “proof of work”. At a mining node 104M, new transactions are added to a pool of valid transactions that have not yet appeared in a block. The miners then race to assemble a new valid block 151 of transactions 152 from the pool of transactions 154 by attempting to solve a cryptographic puzzle. Typically this comprises searching for a “nonce” value such that when the nonce is concatenated with the pool of transactions 154 and hashed, then the output of the hash meets a predetermined condition. E.g. the predetermined condition may be that the output of the hash has a certain predefined number of leading zeros. A property of a hash function is that it has an unpredictable output with respect to its input. Therefore this search can only be performed by brute force, thus consuming a substantive amount of processing resource at each node 104M that is trying to solve the puzzle.

The first miner node 104M to solve the puzzle announces this to the network 106, providing the solution as proof which can then be easily checked by the other nodes 104 in the network (once given the solution to a hash it is straightforward to check that it causes the output of the hash to meet the condition). The pool of transactions 154 for which the winner solved the puzzle then becomes recorded as a new block 151 in the blockchain 150 by at least some of the nodes 104 acting as storage nodes 104S, based on having checked the winner’s announced solution at each such node. A block pointer 155 is also assigned to the new block 151n pointing back to the previously created block 151n-1 in the chain. The proof-of-work helps reduce the risk of double spending since it takes a large amount of effort to create a new block 151, and as any block containing a double spend is likely to be

rejected by other nodes 104, mining nodes 104M are incentivised not to allow double spends to be included in their blocks. Once created, the block 151 cannot be modified since it is recognized and maintained at each of the storing nodes 104S in the P2P network 106 according to the same protocol. The block pointer 155 also imposes a sequential order to the blocks 151. Since the transactions 152 are recorded in the ordered blocks at each storage node 104S in a P2P network 106, this therefore provides an immutable public ledger of the transactions.

Note that different miners 104M racing to solve the puzzle at any given time may be doing so based on different snapshots of the unmined transaction pool 154 at any given time, depending on when they started searching for a solution. Whoever solves their respective puzzle first defines which transactions 152 are included in the next new block 151n, and the current pool 154 of unmined transactions is updated. The miners 104M then continue to race to create a block from the newly defined outstanding pool 154, and so forth. A protocol also exists for resolving any “fork” that may arise, which is where two miners 104M solve their puzzle within a very short time of one another such that a conflicting view of the blockchain gets propagated. In short, whichever prong of the fork grows the longest becomes the definitive blockchain 150.

In most blockchains the winning miner 104M is automatically rewarded with a special kind of new transaction which creates a new quantity of the digital asset out of nowhere (as opposed to normal transactions which transfer an amount of the digital asset from one user to another). Hence the winning node is said to have “mined” a quantity of the digital asset. This special type of transaction is sometime referred to as a “generation” transaction. It automatically forms part of the new block 151n. This reward gives an incentive for the miners 104M to participate in the proof-of-work race. Often a regular (non-generation) transaction 152 will also specify an additional transaction fee in one of its outputs, to further reward the winning miner 104M that created the block 151n in which that transaction was included.

Due to the computational resource involved in mining, typically at least each of the miner nodes 104M takes the form of a server comprising one or more physical server units, or even whole a data centre. Each forwarding node 104M and/or storage node 104S may also take the form of a server or data centre. However in principle any given node 104 could take
5 the form of a user terminal or a group of user terminals networked together.

The memory of each node 104 stores software 400 configured to run on the processing apparatus of the node 104 in order to perform its respective role or roles and handle transactions 152 in accordance with the node protocol. It will be understood that any action
10 attributed herein to a node 104 may be performed by the software 400 run on the processing apparatus of the respective computer equipment. The node software 400 may be implemented in one or more applications at the application layer, or a lower layer such as the operating system layer or a protocol layer, or any combination of these. Also, the term “blockchain” as used herein is a generic term that refers to the kind of technology in
15 general, and does not limit to any particular proprietary blockchain, protocol or service.

Also connected to the network 101 is the computer equipment 102 of each of a plurality of parties 103 in the role of consuming users. These act as payers and payees in transactions but do not necessarily participate in mining or propagating transactions on behalf of other
20 parties. They do not necessarily run the mining protocol. Two parties 103 and their respective equipment 102 are shown for illustrative purposes: a first party 103a and his/her respective computer equipment 102a, and a second party 103b and his/her respective computer equipment 102b. It will be understood that many more such parties 103 and their respective computer equipment 102 may be present and participating in the system, but for
25 convenience they are not illustrated. Each party 103 may be an individual or an organization. Purely by way of illustration the first party 103a is referred to herein as Alice and the second party 103b is referred to as Bob, but it will be appreciated that this is not limiting and any reference herein to Alice or Bob may be replaced with “first party” and “second “party” respectively.

The computer equipment 102 of each party 103 comprises respective processing apparatus comprising one or more processors, e.g. one or more CPUs, GPUs, other accelerator processors, application specific processors, and/or FPGAs. The computer equipment 102 of each party 103 further comprises memory, i.e. computer-readable storage in the form of a non-transitory computer-readable medium or media. This memory may comprise one or more memory units employing one or more memory media, e.g. a magnetic medium such as hard disk; an electronic medium such as an SSD, flash memory or EEPROM; and/or an optical medium such as an optical disc drive. The memory on the computer equipment 102 of each party 103 stores software comprising a respective instance of at least one client application 105 arranged to run on the processing apparatus. It will be understood that any action attributed herein to a given party 103 may be performed using the software run on the processing apparatus of the respective computer equipment 102. The computer equipment 102 of each party 103 comprises at least one user terminal, e.g. a desktop or laptop computer, a tablet, a smartphone, or a wearable device such as a smartwatch. The computer equipment 102 of a given party 103 may also comprise one or more other networked resources, such as cloud computing resources accessed via the user terminal.

The client application 105 may be initially provided to the computer equipment 102 of any given party 103 on suitable computer-readable storage medium or media, e.g. downloaded from a server, or provided on a removable storage device such as a removable SSD, flash memory key, removable EEPROM, removable magnetic disk drive, magnetic floppy disk or tape, optical disk such as a CD or DVD ROM, or a removable optical drive, etc.

The client application 105 comprises at least a “wallet” function. This has two main functionalities. One of these is to enable the respective user party 103 to create, sign and send transactions 152 to be propagated throughout the network of nodes 104 and thereby included in the blockchain 150. The other is to report back to the respective party the amount of the digital asset that he or she currently owns. In an output-based system, this second functionality comprises collating the amounts defined in the outputs of the various transactions 152 scattered throughout the blockchain 150 that belong to the party in question.

Note: whilst the various client functionality may be described as being integrated into a given client application 105, this is not necessarily limiting and instead any client functionality described herein may instead be implemented in a suite of two or more distinct applications, e.g. interfacing via an API, or one being a plug-in to the other. More generally the client functionality could be implemented at the application layer or a lower layer such as the operating system, or any combination of these. The following will be described in terms of a client application 105 but it will be appreciated that this is not limiting.

10

The instance of the client application or software 105 on each computer equipment 102 is operatively coupled to at least one of the forwarding nodes 104F of the P2P network 106. This enables the wallet function of the client 105 to send transactions 152 to the network 106. The client 105 is also able to contact one, some or all of the storage nodes 104 in order to query the blockchain 150 for any transactions of which the respective party 103 is the recipient (or indeed inspect other parties' transactions in the blockchain 150, since in embodiments the blockchain 150 is a public facility which provides trust in transactions in part through its public visibility). The wallet function on each computer equipment 102 is configured to formulate and send transactions 152 according to a transaction protocol. Each node 104 runs software 400 configured to validate transactions 152 according to a node protocol, and in the case of the forwarding nodes 104F to forward transactions 152 in order to propagate them throughout the network 106. The transaction protocol and node protocol correspond to one another, and a given transaction protocol goes with a given node protocol, together implementing a given transaction model. The same transaction protocol is used for all transactions 152 in the blockchain 150 (though the transaction protocol may allow different subtypes of transaction within it). The same node protocol is used by all the nodes 104 in the network 106 (though it may handle different subtypes of transaction differently in accordance with the rules defined for that subtype, and also different nodes may take on different roles and hence implement different corresponding aspects of the protocol).

30

As mentioned, the blockchain 150 comprises a chain of blocks 151, wherein each block 151 comprises a set of one or more transactions 152 that have been created by a proof-of-work process as discussed previously. Each block 151 also comprises a block pointer 155 pointing back to the previously created block 151 in the chain so as to define a sequential order to the blocks 151. The blockchain 150 also comprises a pool of valid transactions 154 waiting to be included in a new block by the proof-of-work process. Each transaction 152 comprises a pointer back to a previous transaction so as to define an order to sequences of transactions (N.B. sequences of transactions 152 are allowed to branch). The chain of blocks 151 goes all the way back to a genesis block (Gb) 153 which was the first block in the chain. One or more original transactions 152 early on in the chain 150 pointed to the genesis block 153 rather than a preceding transaction.

When a given party 103, say Alice, wishes to send a new transaction 152j to be included in the blockchain 150, then she formulates the new transaction in accordance with the relevant transaction protocol (using the wallet function in her client application 105). She then sends the transaction 152 from the client application 105 to one of the one or more forwarding nodes 104F to which she is connected. E.g. this could be the forwarding node 104F that is nearest or best connected to Alice's computer 102. When any given node 104 receives a new transaction 152j, it handles it in accordance with the node protocol and its respective role. This comprises first checking whether the newly received transaction 152j meets a certain condition for being "valid", examples of which will be discussed in more detail shortly. In some transaction protocols, the condition for validation may be configurable on a per-transaction basis by scripts included in the transactions 152. Alternatively the condition could simply be a built-in feature of the node protocol, or be defined by a combination of the script and the node protocol.

On condition that the newly received transaction 152j passes the test for being deemed valid (i.e. on condition that it is "validated"), any storage node 104S that receives the transaction 152j will add the new validated transaction 152 to the pool 154 in the copy of the blockchain 150 maintained at that node 104S. Further, any forwarding node 104F that receives the transaction 152j will propagate the validated transaction 152 onward to one or

more other nodes 104 in the P2P network 106. Since each forwarding node 104F applies the same protocol, then assuming the transaction 152j is valid, this means it will soon be propagated throughout the whole P2P network 106.

- 5 Once admitted to the pool 154 in the copy of the blockchain 150 maintained at one or more storage nodes 104, then miner nodes 104M will start competing to solve the proof-of-work puzzle on the latest version of the pool 154 including the new transaction 152 (other miners 104M may still be trying to solve the puzzle based on the old view of the pool 154, but whoever gets there first will define where the next new block 151 ends and the new pool
- 10 154 starts, and eventually someone will solve the puzzle for a part of the pool 154 which includes Alice's transaction 152j). Once the proof-of-work has been done for the pool 154 including the new transaction 152j, it immutably becomes part of one of the blocks 151 in the blockchain 150. Each transaction 152 comprises a pointer back to an earlier transaction, so the order of the transactions is also immutably recorded.

15

- Different nodes 104 may receive different instances of a given transaction first and therefore have conflicting views of which instance is 'valid' before one instance is mined into a block 150, at which point all nodes 104 agree that the mined instance is the only valid instance. If a node 104 accepts one instance as valid, and then discovers that a second
- 20 instance has been recorded in the blockchain 150 then that node 104 must accept this and will discard (i.e. treat as invalid) the unmined instance which it had initially accepted.

UTXO-BASED MODEL

- 25 Figure 2 illustrates an example transaction protocol. This is an example of an UTXO-based protocol. A transaction 152 (abbreviated "Tx") is the fundamental data structure of the blockchain 150 (each block 151 comprising one or more transactions 152). The following will be described by reference to an output-based or "UTXO" based protocol. However, this not limiting to all possible embodiments.

30

In a UTXO-based model, each transaction ("Tx") 152 comprises a data structure comprising one or more inputs 202, and one or more outputs 203. Each output 203 may comprise an unspent transaction output (UTXO), which can be used as the source for the input 202 of another new transaction (if the UTXO has not already been redeemed). The UTXO specifies
5 an amount of a digital asset (a store of value). It may also contain the transaction ID of the transaction from which it came, amongst other information. The transaction data structure may also comprise a header 201, which may comprise an indicator of the size of the input field(s) 202 and output field(s) 203. The header 201 may also include an ID of the transaction. In embodiments the transaction ID is the hash of the transaction data
10 (excluding the transaction ID itself) and stored in the header 201 of the raw transaction 152 submitted to the miners 104M.

Say Alice 103a wishes to create a transaction 152j transferring an amount of the digital asset in question to Bob 103b. In Figure 2 Alice's new transaction 152j is labelled " T_{X1} ". It takes an
15 amount of the digital asset that is locked to Alice in the output 203 of a preceding transaction 152i in the sequence, and transfers at least some of this to Bob. The preceding transaction 152i is labelled " T_{X0} " in Figure 2. T_{X0} and T_{X1} are just an arbitrary labels. They do not necessarily mean that T_{X0} is the first transaction in the blockchain 151, nor that T_{X1} is the immediate next transaction in the pool 154. T_{X1} could point back to any preceding
20 (i.e. antecedent) transaction that still has an unspent output 203 locked to Alice.

The preceding transaction T_{X0} may already have been validated and included in the blockchain 150 at the time when Alice creates her new transaction T_{X1} , or at least by the time she sends it to the network 106. It may already have been included in one of the blocks
25 151 at that time, or it may be still waiting in the pool 154 in which case it will soon be included in a new block 151. Alternatively T_{X0} and T_{X1} could be created and sent to the network 102 together, or T_{X0} could even be sent after T_{X1} if the node protocol allows for buffering "orphan" transactions. The terms "preceding" and "subsequent" as used herein in the context of the sequence of transactions refer to the order of the transactions in the
30 sequence as defined by the transaction pointers specified in the transactions (which transaction points back to which other transaction, and so forth). They could equally be

replaced with “predecessor” and “successor”, or “antecedent” and “descendant”, “parent” and “child”, or such like. It does not necessarily imply an order in which they are created, sent to the network 106, or arrive at any given node 104. Nevertheless, a subsequent transaction (the descendent transaction or “child”) which points to a preceding transaction (the antecedent transaction or “parent”) will not be validated until and unless the parent transaction is validated. A child that arrives at a node 104 before its parent is considered an orphan. It may be discarded or buffered for a certain time to wait for the parent, depending on the node protocol and/or miner behaviour.

- 10 One of the one or more outputs 203 of the preceding transaction Tx_0 comprises a particular UTXO, labelled here $UTXO_0$. Each UTXO comprises a value specifying an amount of the digital asset represented by the UTXO, and a locking script which defines a condition which must be met by an unlocking script in the input 202 of a subsequent transaction in order for the subsequent transaction to be validated, and therefore for the UTXO to be successfully
- 15 redeemed. Typically the locking script locks the amount to a particular party (the beneficiary of the transaction in which it is included). I.e. the locking script defines an unlocking condition, typically comprising a condition that the unlocking script in the input of the subsequent transaction comprises the cryptographic signature of the party to whom the preceding transaction is locked.

20

The locking script (aka scriptPubKey) is a piece of code written in the domain specific language recognized by the node protocol. A particular example of such a language is called “Script” (capital S). The locking script specifies what information is required to spend a transaction output 203, for example the requirement of Alice’s signature. Unlocking scripts appear in the outputs of transactions. The unlocking script (aka scriptSig) is a piece of code written the domain specific language that provides the information required to satisfy the locking script criteria. For example, it may contain Bob’s signature. Unlocking scripts appear in the input 202 of transactions.

- 25
- 30 So in the example illustrated, $UTXO_0$ in the output 203 of Tx_0 comprises a locking script [Checksig P_A] which requires a signature Sig P_A of Alice in order for $UTXO_0$ to be redeemed

(strictly, in order for a subsequent transaction attempting to redeem $UTXO_0$ to be valid).
 [Checksig P_A] contains the public key P_A from a public-private key pair of Alice. The input
 202 of TX_1 comprises a pointer pointing back to TX_0 (e.g. by means of its transaction ID,
 $TXID_0$, which in embodiments is the hash of the whole transaction TX_0). The input 202 of
 5 TX_1 comprises an index identifying $UTXO_0$ within TX_0 , to identify it amongst any other
 possible outputs of TX_0 . The input 202 of TX_1 further comprises an unlocking script $\langle \text{Sig } P_A \rangle$
 which comprises a cryptographic signature of Alice, created by Alice applying her private key
 from the key pair to a predefined portion of data (sometimes called the “message” in
 cryptography). What data (or “message”) needs to be signed by Alice to provide a valid
 10 signature may be defined by the locking script, or by the node protocol, or by a combination
 of these.

When the new transaction TX_1 arrives at a node 104, the node applies the node protocol.
 This comprises running the locking script and unlocking script together to check whether the
 15 unlocking script meets the condition defined in the locking script (where this condition may
 comprise one or more criteria). In embodiments this involves concatenating the two scripts:

$$\langle \text{Sig } P_A \rangle \langle P_A \rangle || [\text{Checksig } P_A]$$

20 where “||” represents a concatenation and “ $\langle \dots \rangle$ ” means place the data on the stack, and
 “[...]” is a function comprised by the unlocking script (in this example a stack-based
 language). Equivalently the scripts may be run one after the other, with a common stack,
 rather than concatenating the scripts. Either way, when run together, the scripts use the
 public key P_A of Alice, as included in the locking script in the output of TX_0 , to authenticate
 25 that the locking script in the input of TX_1 contains the signature of Alice signing the expected
 portion of data. The expected portion of data itself (the “message”) also needs to be
 included in TX_0 order to perform this authentication. In embodiments the signed data
 comprises the whole of TX_0 (so a separate element does not need to be included specifying
 the signed portion of data in the clear, as it is already inherently present).

The details of authentication by public-private cryptography will be familiar to a person skilled in the art. Basically, if Alice has signed a message by encrypting it with her private key, then given Alice's public key and the message in the clear (the unencrypted message), another entity such as a node 104 is able to authenticate that the encrypted version of the message must have been signed by Alice. Signing typically comprises hashing the message, signing the hash, and tagging this onto the clear version of the message as a signature, thus enabling any holder of the public key to authenticate the signature. Note therefore that any reference herein to signing a particular piece of data or part of a transaction, or such like, can in embodiments mean signing a hash of that piece of data or part of the transaction.

10

If the unlocking script in TX_I meets the one or more conditions specified in the locking script of TX_O (so in the example shown, if Alice's signature is provided in TX_I and authenticated), then the node 104 deems TX_I valid. If it is a storage node 104S, this means it will add it to the pool of transactions 154 awaiting proof-of-work. If it is a forwarding node 104F, it will forward the transaction TX_I to one or more other nodes 104 in the network 106, so that it will be propagated throughout the network. Once TX_I has been validated and included in the blockchain 150, this defines $UTXO_O$ from TX_O as spent. Note that TX_I can only be valid if it spends an unspent transaction output 203. If it attempts to spend an output that has already been spent by another transaction 152, then TX_I will be invalid even if all the other conditions are met. Hence the node 104 also needs to check whether the referenced UTXO in the preceding transaction TX_O is already spent (has already formed a valid input to another valid transaction). This is one reason why it is important for the blockchain 150 to impose a defined order on the transactions 152. In practice a given node 104 may maintain a separate database marking which UTXOs 203 in which transactions 152 have been spent, but ultimately what defines whether a UTXO has been spent is whether it has already formed a valid input to another valid transaction in the blockchain 150.

If the total amount specified in all the outputs 203 of a given transaction 152 is greater than the total amount pointed to by all its inputs 202, this is another basis for invalidity in most transaction models. Therefore such transactions will not be propagated nor mined into blocks 151.

30

- Note that in UTXO-based transaction models, a given UTXO needs to be spent as a whole. It cannot “leave behind” a fraction of the amount defined in the UTXO as spent while another fraction is spent. However the amount from the UTXO can be split between multiple
- 5 outputs of the next transaction. E.g. the amount defined in $UTXO_0$ in T_{X_0} can be split between multiple UTXOs in T_{X_1} . Hence if Alice does not want to give Bob all of the amount defined in $UTXO_0$, she can use the remainder to give herself change in a second output of T_{X_1} , or pay another party.
- 10 In practice Alice will also usually need to include a fee for the winning miner, because nowadays the reward of the generation transaction alone is not typically sufficient to motivate mining. If Alice does not include a fee for the miner, T_{X_0} will likely be rejected by the miner nodes 104M, and hence although technically valid, it will still not be propagated and included in the blockchain 150 (the miner protocol does not force miners 104M to
- 15 accept transactions 152 if they don’t want). In some protocols, the mining fee does not require its own separate output 203 (i.e. does not need a separate UTXO). Instead any difference between the total amount pointed to by the input(s) 202 and the total amount of specified in the output(s) 203 of a given transaction 152 is automatically given to the winning miner 104. E.g. say a pointer to $UTXO_0$ is the only input to T_{X_1} , and T_{X_1} has only
- 20 one output $UTXO_1$. If the amount of the digital asset specified in $UTXO_0$ is greater than the amount specified in $UTXO_1$, then the difference automatically goes to the winning miner 104M. Alternatively or additionally however, it is not necessarily excluded that a miner fee could be specified explicitly in its own one of the UTXOs 203 of the transaction 152.
- 25 Alice and Bob’s digital assets consist of the unspent UTXOs locked to them in any transactions 152 anywhere in the blockchain 150. Hence typically, the assets of a given party 103 are scattered throughout the UTXOs of various transactions 152 throughout the blockchain 150. There is no one number stored anywhere in the blockchain 150 that defines the total balance of a given party 103. It is the role of the wallet function in the client
- 30 application 105 to collate together the values of all the various UTXOs which are locked to the respective party and have not yet been spent in another onward transaction. It can do

this by querying the copy of the blockchain 150 as stored at any of the storage nodes 104S, e.g. the storage node 104S that is closest or best connected to the respective party's computer equipment 102.

- 5 Note that the script code is often represented schematically (i.e. not the exact language). For example, one may write $[\text{Checksig } P_A]$ to mean $[\text{Checksig } P_A] = \text{OP_DUP OP_HASH160 } \langle H(P_A) \rangle \text{ OP_EQUALVERIFY OP_CHECKSIG}$. "OP_..." refers to a particular opcode of the Script language. OP_CHECKSIG (also called "Checksig") is a Script opcode that takes two inputs (signature and public key) and verifies the signature's validity using the Elliptic Curve Digital
- 10 Signature Algorithm (ECDSA). At runtime, any occurrences of signature ('sig') are removed from the script but additional requirements, such as a hash puzzle, remain in the transaction verified by the 'sig' input. As another example, OP_RETURN is an opcode of the Script language for creating an unspendable output of a transaction that can store metadata within the transaction, and thereby record the metadata immutably in the blockchain 150.
- 15 E.g. the metadata could comprise a document which it is desired to store in the blockchain.

Note: the notation $\langle H(x) \rangle$ means "push the value h onto the stack", where the value $h = H(x)$ is provided in the unlocking script, without providing H or x .

- 20 The signature P_A is a digital signature. In embodiments this is based on the ECDSA using the elliptic curve secp256k1. A digital signature signs a particular piece of data. In embodiments, for a given transaction the signature will sign part of the transaction input, and all or part of the transaction output. The particular parts of the outputs it signs depends on the SIGHASH flag. The SIGHASH flag is a 4-byte code included at the end of a signature to select which
- 25 outputs are signed (and thus fixed at the time of signing).

- The locking script is sometimes called "scriptPubKey" referring to the fact that it comprises the public key of the party to whom the respective transaction is locked. The unlocking script is sometimes called "scriptSig" referring to the fact that it supplies the corresponding
- 30 signature. However, more generally it is not essential in all applications of a blockchain 150 that the condition for a UTXO to be redeemed comprises authenticating a signature. More

generally the scripting language could be used to define any one or more conditions. Hence the more general terms “locking script” and “unlocking script” may be preferred.

OPTIONAL SIDE CHANNEL

5

Figure 3 shows a further system 100 for implementing a blockchain 150. The system 100 is substantially the same as that described in relation to Figure 1 except that additional communication functionality is involved. The client application on each of Alice and Bob’s computer equipment 102a, 120b, respectively, comprises additional communication
10 functionality. That is, it enables Alice 103a to establish a separate side channel 301 with Bob 103b (at the instigation of either party or a third party). The side channel 301 enables exchange of data separately from the P2P network. Such communication is sometimes referred to as “off-chain”. For instance this may be used to exchange a transaction 152 between Alice and Bob without the transaction (yet) being published onto the network P2P
15 106 or making its way onto the chain 150, until one of the parties chooses to broadcast it to the network 106. Alternatively or additionally, the side channel 301 may be used to exchange any other transaction related data, such as keys, negotiated amounts or terms, data content, etc.

20 The side channel 301 may be established via the same packet-switched network 101 as the P2P overlay network 106. Alternatively or additionally, the side channel 301 may be established via a different network such as a mobile cellular network, or a local area network such as a local wireless network, or even a direct wired or wireless link between Alice and Bob’s devices 1021, 102b. Generally, the side channel 301 as referred to anywhere
25 herein may comprise any one or more links via one or more networking technologies or communication media for exchanging data “off-chain”, i.e. separately from the P2P overlay network 106. Where more than one link is used, then the bundle or collection of off-chain links as a whole may be referred to as the side channel 301. Note therefore that if it is said that Alice and Bob exchange certain pieces of information or data, or such like, over the side
30 channel 301, then this does not necessarily imply all these pieces of data have to be send over exactly the same link or even the same type of network.

NODE SOFTWARE

Figure 4 illustrates an example of the node software 400 that is run on each node 104 of the P2P network 106, in the example of a UTXO- or output-based model. The node software 400 comprises a protocol engine 401, a script engine 402, a stack 403, an application-level decision engine 404, and a set of one or more blockchain-related functional modules 405. At any given node 104, these may include any one, two or all three of: a mining module 405M, a forwarding module 405F and a storing module 405S (depending on the role or roles of the node). The protocol engine 401 is configured to recognize the different fields of a transaction 152 and process them in accordance with the node protocol. When a transaction 152m (Tx_m) is received having an input pointing to an output (e.g. UTXO) of another, preceding transaction 152m-1 (Tx_{m-1}), then the protocol engine 401 identifies the unlocking script in Tx_m and passes it to the script engine 402. The protocol engine 401 also identifies and retrieves Tx_{m-1} based on the pointer in the input of Tx_m . It may retrieve Tx_{m-1} from the respective node's own pool 154 of pending transactions if Tx_{m-1} is not already on the blockchain 150, or from a copy of a block 151 in the blockchain 150 stored at the respective node or another node 104 if Tx_{m-1} is already on the blockchain 150. Either way, the script engine 401 identifies the locking script in the pointed-to output of Tx_{m-1} and passes this to the script engine 402.

The script engine 402 thus has the locking script of Tx_{m-1} and the unlocking script from the corresponding input of Tx_m . For example Tx_1 and Tx_2 are illustrated in Figure 4, but the same could apply for any pair of transactions, such as Tx_0 and Tx_1 , etc. The script engine 402 runs the two scripts together as discussed previously, which will include placing data onto and retrieving data from the stack 403 in accordance with the stack-based scripting language being used (e.g. Script).

By running the scripts together, the script engine 402 determines whether or not the unlocking script meets the one or more criteria defined in the locking script – i.e. does it “unlock” the output in which the locking script is included? The script engine 402 returns a

result of this determination to the protocol engine 401. If the script engine 402 determines that the unlocking script does meet the one or more criteria specified in the corresponding locking script, then it returns the result “true”. Otherwise it returns the result “false”.

- 5 In an output-based model, the result “true” from the script engine 402 is one of the conditions for validity of the transaction. Typically there are also one or more further, protocol-level conditions evaluated by the protocol engine 401 that must be met as well; such as that the total amount of digital asset specified in the output(s) of Tx_m does not exceed the total amount pointed to by its inputs, and that the pointed-to output of
- 10 Tx_{m-1} has not already been spent by another valid transaction. The protocol engine 401 evaluates the result from the script engine 402 together with the one or more protocol-level conditions, and only if they are all true does it validate the transaction Tx_m . The protocol engine 401 outputs an indication of whether the transaction is valid to the application-level decision engine 404. Only on condition that Tx_m is indeed validated, the decision engine
- 15 404 may select to control one or both of the mining module 405M and the forwarding module 405F to perform their respective blockchain-related function in respect of Tx_m . This may comprise the mining module 405M adding Tx_m to the node’s respective pool 154 for mining into a block 151, and/or the forwarding module 405F forwarding Tx_m to another node 104 in the P2P network 106. Note however that in embodiments, while the decision
- 20 engine 404 will not select to forward or mine an invalid transaction, this does not necessarily mean that, conversely, it is obliged to trigger the mining or the forwarding of a valid transaction simply because it is valid. Optionally, in embodiments the decision engine 404 may apply one or more additional conditions before triggering either or both of these functions. E.g. if the node is a mining node 104M, the decision engine may only select to
- 25 mine the transaction on condition that the transaction is both valid and leaves enough of a mining fee.

- Note also that the terms “true” and “false” herein do not necessarily limit to returning a result represented in the form of only a single binary digit (bit), though that is certainly one
- 30 possible implementation. More generally, “true” can refer to any state indicative of a successful or affirmative outcome, and “false” can refer to any state indicative of an

unsuccessful or non-affirmative outcome. For instance in an account-based model (not illustrated in Figure 4), a result of “true” could be indicated by a combination of an implicit, protocol-level) validation of a signature by the node 104 and an additional affirmative output of a smart contract (the overall result being deemed to signal true if both individual
5 outcomes are true).

EXAMPLE TRANSACTION SET

Figure 5 illustrates a set of transactions 152 for use in accordance with embodiments disclosed herein. The set comprises: a zeroth transaction TX_0 , a first transaction TX_1 , and a second transaction TX_2 . Note that “zeroth”, “first” and “second” are just convenient labels. They do not necessarily imply that these transactions will be placed immediately one after another in a block 151 or the blockchain 150, nor that the zeroth transaction is the initial transaction in a block 151 or the blockchain 150. Nor do these labels necessarily imply
10 anything about the order their transactions are sent to the network 106. They refer only to a logical series in that the output of one transaction is pointed to by the input of the next transaction. Remember that in some systems it is possible to send a parent to the network 106 after its child (in which case the “orphan” child will be buffered for a period at one or more nodes 104 while waiting for the parent to arrive).
15

20 The zeroth transaction TX_0 may also be referred to as the source transaction for the present purposes, in that it acts as a source of an amount of the digital asset which is locked to Alice 103a. The first transaction TX_1 may also be referred to as the challenge transaction or puzzle transaction for the present purposes. It acts as an intermediary for conditionally transferring
25 the amount of digital asset from the source transaction TX_0 in dependence on the second transaction TX_2 providing a solution to the r-puzzle. The second transaction TX_2 may also be referred to as the proving transaction, or spending transaction, as it is the transaction that will provide the solution to the r-puzzle set in the first transaction TX_1 and lock the resulting payment to the prover (or potentially a beneficiary on behalf of whom the prover is acting).
30 Embodiments may be described by way of example whereby the prover (second party) happens to be Bob, but as will be appreciated based on the discussion later, the r-puzzle in

fact allows any second party to be the prover regardless of identity as long as they provide a valid signature solving the r-puzzle.

As shown in Figure 5, the source transaction T_{X0} comprises at least one output 203₀ (e.g.

5 output 0 of T_{X0}) which specifies an amount of the digital asset, and which further comprises a locking script locking this output to Alice 103a. This means that the locking script of the source transaction T_{X0} requires at least one condition to be met, which is that the input of any transaction attempting to unlock the output (and therefore redeem the amount of the digital asset) must include a cryptographic signature of Alice (i.e. using Alice's public key) in
10 its unlocking script. In this sense the amount defined in the output of T_{X0} may be said to be owned by Alice. The output may be referred to as a UTXO. It is not particularly material for the present purposes which output of which preceding transaction the inputs of T_{X0} point back to (as long as they are sufficient to cover the total output(s) of T_{X0}).

15 In the present case the transaction unlocking the output of the source transaction T_{X0} is the first transaction T_{X1} (the challenge transaction). Therefore T_{X1} has at least one input 202₁ (e.g. input 0 of T_{X1}) which comprises a pointer to the relevant output of T_{X0} (output 0 of T_{X0} in the illustrated example), and which further comprises an unlocking script configured to unlock the pointed-to output of T_{X0} according to the condition defined in the locking script
20 of that output, which requires at least a signature of Alice. The signature required from Alice by the locking script of T_{X0} is required to sign some part of T_{X1} . In some protocols the part of T_{X1} that needs to be signed can be a setting defined in the unlocking script of T_{X1} . E.g. this may be set by the SIGHASH flag, which is one byte that is appended to the signature, so in terms of data the unlocking script appears as: <Sig P_A ><sighashflag> < P_A >. Alternatively
25 the part that needs to be signed could simply be a fixed or default part of T_{X1} . Either way, the part to be signed typically excludes the unlocking script itself, and may exclude some or all of the inputs of T_{X1} . The signed part of T_{X1} will however include at least the output 203₁ containing the r-puzzle (see below, Output 0 of T_{X1} in this example).

30 The first transaction T_{X1} has at least one output 203₁ (e.g. output 0 of T_{X1} , which again the output may be referred to as a UTXO). The output of the first transaction T_{X1} is not locked

to any one party. Like T_{X0} it has at least one output (e.g. output 0 of T_{X1}) which specifies an amount of digital asset to be transferred onwards, and which further comprises a locking script defining what is required to unlock that output and hence redeem this amount. However, this locking script allows its output to be unlocked by any party providing a solution to the r-puzzle.

The second transaction (spending transaction) T_{X2} has at least one input 202₂ (e.g. input 0 of T_{X2}) which comprises a pointer to the above-mentioned output of T_{X1} (Output 0 of T_{X1} , in the example shown), and which also comprises an unlocking script configured to unlock said output of T_{X1} based on meeting the one or more requirements of the unlocking condition defined in the locking script of T_{X1} . In accordance with embodiments disclosed herein, the unlocking condition includes at least a requirement that the corresponding unlocking script includes a solution to the r-puzzle. The r-puzzle comprises a challenge defined in the locking script of T_{X1} based on the r-part of an elliptical curve cryptography (ECC) signature, which can be met by any party (in this case happening to be Bob) including their signature (or at least the s-part thereof) in the unlocking script of T_{X2} . Note that unlike the locking script of T_{X0} , any party's signature can be used to unlock the locking condition in T_{X1} , as long as it is a valid signature that meets the r-challenge (i.e. r-puzzle). Examples of this will be discussed in more detail shortly. Bob is simply chosen as an example of the prover or second party here, but the r-puzzle in fact allows any second party to be the prover, e.g. Charlie, Dora, Ezekiel, etc. In some embodiments, the unlocking condition in T_{X1} could also be made conditional on one or more further conditions, e.g. requiring a signature of Alice to be included in the unlocking script of T_{X2} as well.

The second transaction T_{X2} has at least one output 202₂ (e.g. Output 0 of T_{X2}) which specifies an amount of the digital asset to transfer to Bob, and a locking script locking this to Bob (i.e. it would require a further, onward transaction including Bob's signature in the unlocking script to spend). In this sense the output of the target transaction T_{X2} can be said to be owned by Bob. This output may again be referred to as a UTXO.

The part of T_{X2} signed by the prover's signature (e.g. $Sig P_B$ if it is Bob) will include at least this output 203₂, i.e. the output locking the payment to the prover (Output 0 of T_{X2} in this example).

- 5 In embodiments, it is possible that the locking script in the output 203₁ of T_{X1} defines multiple alternative conditions for unlocking the output, e.g. multiple alternative r-puzzles. In this case the unlocking script in the input 202₂ of T_{X2} unlocks the output of T_{X1} if it meets any one of the alternative unlocking conditions.
- 10 The zeroth (i.e. source) transaction T_{X0} may be generated by Alice, the prover (e.g. Bob) or a third party. It will typically require the signature of the preceding party from whom Alice obtained the amount defined in the input of T_{X0} . It may be sent to the network 106 by Alice, Bob, the preceding party, or another third party.
- 15 The first transaction (i.e. challenge transaction) T_{X1} may also be generated by Alice, the prover (e.g. Bob) or a third party. Since in embodiments it requires Alice's signature, it may be generated by Alice. Alternatively it may be generated by Bob or a third party as a template then sent to Alice to sign, e.g. being sent over the side channel 301. Alice can then send the signed transaction to the network 106 herself, or send it to Bob or a third party for
- 20 them to forward to the network 106, or just send her signature for Bob or the third party to assemble into the signed T_{X1} and forward to the network 106. Any off-chain exchanges prior to sending T_{X1} to the network 106 may be performed over the side channel 301.

- The second transaction (i.e. proving or spending transaction) T_{X2} may be generated by Alice,
- 25 the prover (e.g. Bob) or a third party. As the first version requires the prover's signature and/or data, it may be generated by Bob. Alternatively it may be generated as a template by Alice or a third party then sent to Bob to sign, e.g. being sent to Bob over the side channel 301. Bob can then send the signed transaction to the network 106 himself, or send it to Alice or a third party for them to forward to the network 106, or just send his signature and
- 30 for Alice or the third party to assemble into the signed T_{X2} and forward to the network.

It will be appreciated that there are various locations at which the different elements of a transaction can be generated and assembled, and various ways for it to be sent onwards directly or vicariously to the ultimate destination of the P2P network 106. The scope of implementation of the disclosed techniques is not limited in any of these respects.

5

It will also be appreciated that phrases such as “by Alice”, “by Bob” and “by a third party” herein may be used as a short-hand for “by the computer equipment 102a of Alice 103a”, “by the computer equipment 102b of Bob 103b”, and “by computer equipment of the third party”, respectively. Also, note again that the equipment of a given party could comprise one or more user devices used by that party, or server resources such as cloud resources employed by that party, or any combination of these. It does not necessarily limit the actions to being performed on a single user device.

10

ELLIPTICAL CURVE DIGITAL SIGNATURE ALGORITHMS (ECDSAs)

15

Public key cryptography is used as a basis for securing transactions in a number of different blockchain architectures. Uses of public key cryptography include public key encryption and digital signature schemes. Public key cryptography is founded on the principle that certain functions are easy to compute but hard to reverse without some special knowledge. Such a function is called a trapdoor function and the special knowledge needed to reverse it is referred to as a trapdoor of that function. Easy to compute means it is computationally feasible to compute the trapdoor function for a given input (or set of inputs) in a reasonable time frame, and hard to reverse that it is computationally infeasible to infer that input (or those inputs) from the result without knowledge of the trapdoor.

20

25

In the context of public key cryptography, a key pair means a public key (which can be made freely available to anyone) and a corresponding private key (which is assumed to be secret in the sense that it is only known to a specific entity or group). The public key defines a trapdoor function and the corresponding private key is the trapdoor needed to reverse that function.

30

In a public key encryption context, encryption is based on the trapdoor function (i.e. encryption is performed in the “forward direction”), whereas decryption is based on the reversal the trapdoor function (i.e. decryption is performed in the “reverse direction”) which is only feasible when the trapdoor is known.

5

In a digital signature context, signature verification is performed in the forward direction, using the public key, and signature generation is performed in the reverse direction and can only feasibly be performed using the private key.

- 10 In a blockchain context, digital signatures based on public key cryptography are used as a basis for cryptographically signing transactions and verifying transaction signatures.

ECC is a form of public key cryptography which harnesses the mathematical properties of elliptical curves, and has various benefits over other cryptographic schemes such as DSA

- 15 (Digital Secure Algorithm).

The “Elliptic Curve Digital Signature Algorithm” (ECDSA) refers to a class of digital signature schemes which use ECC as a basis for digital signature generation and verification. Certain principles of the ECDSA are outlined below.

20

In mathematical terminology, ECC exploits the algebraic structure of elliptic curves over finite fields of prime order. A finite field means a finite set of elements and a set of associated operations of multiplication, addition, subtraction and division which satisfy the normal rules of arithmetic (associativity, commutativity etc.) when applied to the elements in the set. That is to say, operations which need not be addition, multiplication etc. in the “normal” sense, but which do behave in essentially the same way.

25

Elliptic Curve Operations:

- 30 In the context of ECC, the addition, subtraction and multiplication operations are, respectively, elliptic curve point addition, denoted “+” herein, elliptic curve point

subtraction, denoted “ $-$ ” herein, and elliptic curve scalar multiplication, denoted “ $\cdot$ ” herein. The addition and subtraction operations are each applied to two points on an elliptic curve and return a third point on the elliptic curve; however, the multiplication operation is applied to a scalar and a single point on an elliptic curve, and returns a second point on the elliptic curve. Division is, by contrast, defined on scalars.

For the purposes of illustration, Figure 6A shows an elliptic curve $\mathcal{E}$ in $\mathbb{R}^2$, $\mathbb{R}^2$ being the set of all real-valued two-dimensional coordinates and $(x, y) \in \mathbb{R}^2$ denoting an element of $\mathbb{R}^2$. The elliptical curve $\mathcal{E}$ is the set of points which satisfy the following equation:

$$\mathcal{E}: y^2 = x^3 + ax + b$$

Addition: A mathematical property of $\mathcal{E}$ is that, given any two points A, B on the elliptic curve $\mathcal{E}$, a line intersecting A and B will re-intersect $\mathcal{E}$ and one additional point only, denoted C ; the elliptic curve addition of A and B , i.e. $A + B$, is defined as the “reflection” of C : taking the horizontal line which intersects C , the reflection of C is the other point on the elliptic curve intersected by that line. This definition hold for the case $A = B$, with the modification that C is now the point at which the tangent to $\mathcal{E}$ at A re-intersects $\mathcal{E}$. This definition is made to hold for the case that the line intersecting two points is vertical by defining a point at infinity, denoted ∞ , as a point on the elliptic curve and at which any vertical line intersects the elliptic curve (e.g. the points labelled D and E are vertically horizontally aligned, hence $D + E = \infty$).

Subtraction/additive inverse: The above definition of reflection applies to any point, and provides the definition of elliptic curve point subtraction: $A - B$ is the sum of A with the reflection of B . The reflection of B is more formally referred to as the “additive inverse” of B , which in turn is denoted $-B$. Using this notation, elliptic curve subtraction can be defined in mathematical notation as:

$$A - B = A + (-B).$$

Hence, in Figure 6B, $C = -(A + B)$ and $(A + B) = -C$. Note also that, under this definition, $D = -E$, which reflects a general rule of the algebraic structure, namely that the elliptic point addition of any point on the elliptic curve with its additive inverse is the point at infinity, i.e.

$$A + (-A) = \infty \quad \forall A \in \mathcal{E}$$

The point at infinity ∞ is more formally referred to as an “identity element” (note both the parallel with and the deviation from normal arithmetic: in normal arithmetic, the sum of any number a with its additive inverse $-a$ is 0, with 0 being the identity element for normal arithmetic). Another property of the identity element, ∞ which mirrors normal arithmetic, is that $A + \infty = A$ for any point A on $\mathcal{E}$ including ∞ itself (analogous to the statement $a + 0 = a$ for any real number a)

Multiplication: From the definition of elliptic curve point addition, the definition of elliptic curve scalar multiplication follows: the multiplication of an elliptic curve point A with an integer v is defined as:

$$v \cdot A = \underbrace{A + \dots + A}_{v \text{ times}}.$$

That is, as v elliptic curve point additions of A with itself.

20

Note: elliptic curve scalar multiplication is also referred to in the art as elliptic curve point multiplication. Those two terms have the same meaning in the present disclosure.

Division/multiplicative Inverse: The operation of division is defined with respect to scalars: given a scalar v , its “multiplicative inverse” is defined at the scalar v^{-1} such that:

25

$$vv^{-1} = 1.$$

Figure 6A provides an intuitive visualization of the above operations, in which $\mathcal{E}$ is defined over an *infinite* field comprising all real-numbers $\mathbb{R}$.

30

Figure 6B more closely represents how the above operations are actually applied in the context of ECC, as it shows an elliptic curve $\mathcal{E}_n$ defined by the equation:

$$\mathcal{E}_n: y^2 = x^3 + ax + b \mod p$$

5

where p is a prime number (the prime modulus) and mod denotes the modulo operation. The set of points which satisfy the above equation is finite, and all but one of those points are represented in Figure 6B as white circles; the remaining point is the identity element ∞ . The prime number p forms part of the definition of the elliptic curve, and can be freely
 10 chosen. For the elliptic curve to have good cryptographic properties, p should be sufficiently large. For example, a 256 bit p is specified in certain blockchain models.

The subscript “ n ”, by contrast, is referred to herein as the order of the group formed by the elliptic curve points under the point addition defined above (as shorthand, this may be
 15 called the order of the elliptic curve $\mathcal{E}_n$) - see below.

In other words, n is the order of the group, and p is the order of the field. There will be n elliptic curve points in total. Each point on the elliptic curve is represented by two numbers/coordinates (x, y) , where x and y are all in the range $-(p - 1), \dots, 0, \dots, (p - 1)$.

20

It can be seen that $\mathcal{E}_n$ in Figure 6B exhibits a horizontal symmetry which is analogous to that of $\mathcal{E}$ in Figure 6A, which is a general property of elliptic curves over prime fields, hence the definition of the additive inverse of a point on $\mathcal{E}_n$ still holds. Some points have no horizontally-aligned counterpoint (e.g. $(0,0)$) and such points are their own additive inverse.

25

The “line” $l_{A,B}$ intersecting two points A and B on $\mathcal{E}_n$ becomes a finite set of points, represented by smaller black circles, satisfying analogous geometric requirements, and the definition of elliptic curve scalar multiplication still holds. Analogous with Figure 6A, Figure 6B shows the point $A + B = -C$, which is the additive inverse of the point $C = -(A + B)$ at
 30 which the line $l_{A,B}$ re-intersects $\mathcal{E}_n$.

The elliptic curve addition $A + B = -C$ of any two points on $\mathcal{E}_n$ can be defined algebraically by the following equations:

$$\begin{aligned}
 A &= (x_A, y_A), \\
 B &= (x_B, y_B), \\
 5 \quad C &= (x_C, y_C) = -(A + B), \\
 x_C &= (\lambda^2 - x_A - x_B) \bmod p, \\
 y_C &= (\lambda(x_C - x_A) + y_A) \bmod p, \\
 &= (\lambda(x_C - x_B) + y_B) \bmod p,
 \end{aligned}$$

where

$$10 \quad \lambda = (y_A - y_B)(x_A - x_B)^{-1} \bmod p \quad \text{if } A \neq B,$$

and

$$\lambda = (2y_A)^{-1}(3x_A^2 + a) \bmod p \quad \text{if } A = B.$$

For the purposes of the above, the definition of the multiplicate inverse v^{-1} of an integer v is modified as:

$$v^{-1}v \equiv 1 \pmod{p}.$$

That is, the multiplicate inverse of the integer v is the modular inverse of $v \bmod p$.

20

The case of $B = -A$ is special, and resolved by the introduction of the identity element ∞ - as noted, in that case $A + B = A + (-A) = \infty$. The case of $B = \infty$ is also a special case, resolved as noted above as $A + \infty = A$.

25 The definition of elliptic curve scalar multiplication adopts this definition of elliptic curve addition and otherwise remains the same.

In other contexts, the definition of the multiplicative inverse v^{-1} of a *scalar* v with respect is:

30

$$v^{-1}v \equiv 1 \pmod{n}$$

It will be clear in context whether a multiplicative inverse is defined with respect to $\text{mod } n$ or $\text{mod } p$.

In practice, to identify whether a number should be treated as $\text{mod } n$ or $\text{mod } p$, the

5 following checks may be applied:

1. Is the number representing a coordinate of an EC point?

a. If yes, then $\text{mod } p$

2. Is the number to be used to multiply an EC point?

a. If yes, then $\text{mod } n$

10 Note that, there are occasions where both checks give positive answer, in which case that the number has to be $\text{mod } p$ and $\text{mod } n$.

Elliptic Curve Cryptography (ECC)

15 Elliptic curve arithmetic provides unique capabilities in obscuring a secret value and forms the basis of many contemporary cryptographic systems. In particular, reversing scalar multiplication of elliptic curve points over finite fields is an intractable problem (it is computationally infeasible to perform).

20 A private key V takes the form of an integer, and the corresponding public key P is a point P on the elliptic curve $\mathcal{E}_n$ derived from a “generator point” G , which is also a point on the elliptic curve $\mathcal{E}_n$, as:

$$P = V \cdot G = \underbrace{G + \dots + G}_{V \text{ times}}$$

25 where ‘ $\cdot$ ’ denotes elliptic curve scalar multiplication on the elliptic curve $\mathcal{E}_n$ defined by a, b and n (the elliptic curve parameters).

The order n of the elliptic curve $\mathcal{E}_n$ is defined, with respect the generator point G , as the scalar satisfying the following equation:

30
$$n \cdot G = \infty ,$$

Hence, n is determined both by the choice of elliptic curve and the choice of generator point G . Although n may be referred to as the order of the elliptic curve $\mathcal{E}_n$ as shorthand, more precisely n is the order of the (subgroup generated by the) generator point G with respect to the elliptic curve. In a practical ECC context, G should be chosen so that n is large and
5 prime. Therefore, in practice, p and n are both prime numbers, but are not equal to each other in general.

For a sufficiently large V , actually performing V elliptic curve additions to derive P is hard, i.e. computationally infeasible. However, if V is known, then P can be computed much
10 more efficiently by exploiting the algebraic properties of the elliptic curve operations. An example of an efficient algorithm that can be used to compute P is the “double and add” algorithm – crucially, this can only be implemented if V is known.

Conversely, if V is not known, then there is no computationally feasible way of deriving V
15 (i.e. reversing the scalar multiplication) even if both G and P are known (this is the so-called “discrete-logarithm problem”). An attacker could attempt to “brute force” P by starting from G and repeatedly performing elliptic curve point additions until he gets to P ; at that point, he would know V to be the number of elliptic curve point additions he had to perform; but that turns out to be computationally infeasible. Hence, V satisfies the
20 requirements of a trapdoor in the above sense.

In ECC, the public key P , generator key G and elliptic curve $\mathcal{E}_n$ are public and assumed to be known, whereas the private key V is secret.

25 **Elliptic Curve Digital Signature Verification Algorithm (ECDSA)**

In a blockchain system, a user or other entity will typically hold a private key V that is used to prove their identity and the corresponding public key P would be calculated by:

30
$$P = V \cdot G$$

The private key V can be used sign a piece of data m ("the message") using the ECDSA.

Further details of the ECDSA may for example be found in the following, which is incorporated herein by reference in its entirety: "RFC 6979 - Deterministic Usage of the Digital Signature Algorithm (DSA) and Elliptic Curve Digital Signature Algorithm (ECDSA)",
 5 Tools.ietf.org, 2019.

Figure 6C shows a schematic functional block diagram of a signature generation function (signature generator_ 600 which generates an ECDSA signature (r, s) for a public key-
 10 private key pair (V, P) . The EDSA signature is a pair of values, referred to herein as the r-part (r) and s-part (s) respectively.

The signature generation is based on the same elliptic curve $\mathcal{E}_n$ and generator point G used to derived the public key P , hence the elliptic curve parameters a, b and n and the
 15 generator point G are shows as inputs to the signature generator 600.

An ephemeral key generator 602 of the signature generator600 generates an "ephemeral" key $k \in [1, n - 1]$, i.e. in the range from 1 to $n - 1$ inclusive.

20 An r-part generator 604 calculates a corresponding public ephemeral key from k as follows:

$$R = k \cdot G$$

and then take the x-coordinate (with $[]_x$ denoting the process of taking the x coordinate of an elliptic curve point) of the calculated point:

25

$$r = [R]_x$$

which is the r-part of the signature.

As s-part generator 606 calculates an s-part of signature (s) using the modular inverse k^{-1} of $k \bmod n$ (i.e. such that $k^{-1}k \equiv 1 \pmod{n}$ – see above) and a hash of the message m , denoted $H(m)$ (truncated if necessary) as follows:

$$s = k^{-1}(H(m) + rV) \bmod n$$

In the present example, the message m comprises data to be included in a transaction 608 (one or more transaction outputs in the present example). This may be referred to as the process of signing the message m , and the message m may be referred to as a signed part of the transaction.

The message m and the signature (r, s) , in turn, form part of the transaction 608. In the present example, the signature (r, s) is included in an input of the transaction 608 as part of an unlocking script.

Figure 6D shows a schematic functional block diagram of a signature verification function (signature verifier) 620 for verifying the transaction 608. The calculations performed by the signature verifier 620 are based on the same elliptic curve $\mathcal{E}_n$ and generator point G which, as noted, are public.

Whilst the signature requires the private key V as input, that is, one requires knowledge of it in order to generate a valid signature, only the signature pair (r, s) , the message m , and the public key P are needed to validate the signature (r, s) . In order to verify the signature, the signature verifier 620 hashes the signed part of the transaction m (applying the same hash function H as used to generate the signature (r, s)). The verification process is then performed using the following calculation:

$$R' = H(m)s^{-1} \cdot G + rs^{-1} \cdot P$$

The signature is valid (i.e. the signature verification will succeed) if and only if $[R']_x = r$, otherwise it is invalid (i.e. the signature verification fails). In the present example, r denotes the r -part of the signature included in the transaction 608.

- 5 The public key P used in the signature verification process could for example be specified in the locking script of a preceding transaction. The signature verification is performed, in that case, using the public key specified in the locking script of the preceding transaction, and the signed part m and the signature (r, s) of the (later) transaction 608 – and will fail unless the signature (r, s) has been generated based the private key V corresponding to the public
- 10 key P specified in the preceding transaction and the signed part m of the later transaction 608. Hence, only the person who holds the private key V can claim the output of the preceding transaction (typically by including their own public key in the output of the later transaction 608), and the signed part m of the later transaction 608 cannot be altered without invalidating the signature (r, s) .

15

R-PUZZLE

- The following describes a new form of knowledge proof based on an ECDSA. By way of illustration, the challenger is a first party Alice who sets up the r -puzzle in a first transaction
- 20 Tx_1 , either by creating and publishing Tx_1 to the P2P blockchain network 106 herself, or by providing the necessary details to a third party for them to assemble into Tx_1 and publish. The verifier (the party who actually runs the proof) is the operator of a node 104 of the network, e.g. a miner. The solution to the r -puzzle is provided by publishing Tx_2 to the network 106. The prover can be any second party as the r -puzzle is not inherently tied to
- 25 identity, but by way of example the below may be described in terms of a scenario where the prover happens to be Bob. The prover may create and publish Tx_2 himself, or provide the necessary details to a third party for them to assemble into Tx_2 and publish

- Cryptographic hash functions provide a means of deterministically obscuring an input where
- 30 a small change in the input leads to an unpredictable change in the output. Conventional hash functions include MD5, RIPEMD-160, SHA-1, and SHA-256 [5], each of which provide

collision resistance (extremely small probability of finding two inputs that produce the same output) and pre-image resistance (given a hash value $h = H(d)$ it's extremely difficult to find the input d).

- 5 A conventional hash-puzzle can be set up as follows. The idea is to set up a first transaction Tx_1 which allows its output to be redeemed by a second transaction Tx_2 on condition that the second transaction Tx_2 includes some specific piece of data in its input.

In a blockchain transaction, the first party (Alice) could naively create a non-standard

- 10 transaction Tx_1 using a hash value h within the locking script as:

OP_HASH160 < h > OP_EQUALVERIFY

where $h = H_{puz}(d)$ and H_{puz} is a hash function used in the puzzle (in the above example,

- 15 according to the locking script this hash function has to be HASH160, but in other implementations another form of hash function could be used). To redeem the UTXO in which this locking script is included will require the hash puzzle solution in the unlocking script of a subsequent transaction. As such, the spending transaction Tx_2 for a second party with address $Addr_Bob$ would be constructed with an unlocking script which need only
- 20 contain d .

$TxID_2$	
Input	Output
0. TxID1 Unlocking script: < d >	0. Address: $Addr_Bob$ Amount: {VALUE}

where $TxID_i$ is the transaction ID of Tx_i . The locking script says: take the data value d from unlocking script in the input of Tx_2 , hash it, and check whether that equals the hash value h

included in the locking script in the output of Tx_1 . Hence the output it is unlocked by providing d in the unlocking script of Tx_2 .

In this naïve example, after having seen the user's transaction with the hash puzzle solution in Tx_2 , the miner who first receives this transaction can maliciously reject the transaction and create a new, malleated version Tx_2^* with the same solution to the hash puzzle, but changing the output to their own address $Addr_Miner$. The malicious miner can then try to mine Tx_2^* into a block 151 him/herself and if they succeed in mining it before Tx_2 gets mined then the malicious miner will receive the payment instead of Bob.

10

$TxID_2^*$	
Input	Output
1. TxID1 Unlocking script: $\langle d \rangle$	1. Address: $Addr_Miner$ Amount: $\{VALUE\}$

Digital signatures are commonly used in blockchain transactions to prove ownership and redeem unspent transaction outputs (UTXOs). This enables an output of a transaction such as Tx_1 to be locked to a specific party. The most common example is a pay-to-public-key-hash (P2PKH) transaction where the output of the transaction is locked to a particular hash of a public key (which also acts as the address of that party). The locking script for a public key P is:

15

OP_DUP OP_HASH160 $\langle h_P \rangle$ OP_EQUALVERIFY OP_CHECKSIG

20

where $h_P = H_{sig}(P)$ and H_{sig} is a hash function used in the signature (in the above example, according to the locking script this hash function has to be HASH160, but in other implementations another form of hash function could be used). In order to be able to use this UTXO as an input to another transaction, one would have to provide an unlocking script with a valid ECDSA signature using P :

25

$$\langle sig \rangle \langle P \rangle$$

The entire string (unlocking + locking script) is evaluated by the miner, which checks that the correct public key is provided and that the signature is valid and corresponds to P . The locking script basically says: take the public key P from the unlocking script in the input of Tx_2 , hash it, and check whether that equals the hash value h_p included in the locking script in the output of Tx_1 ; and also verify the signature sig using the public key P from the unlocking script of Tx_2 based on an ECDSA verification function, given knowledge of the signed part of Tx_2 . The ECDSA verification function is invoked by the OP_CHECKSIG opcode.

Hence the output can only be unlocked by providing, in the unlocking script of Tx_2 , a valid signature sig signed based on the private key V corresponding to P .

Putting this together with the hash puzzle, the above-mentioned vulnerability can be rectified by requiring a digital signature from the intended recipient, along with the hash puzzle solution. The locking script would be constructed as:

OP_HASH160 $\langle h \rangle$ OP_EQUALVERIFY OP_DUP OP_HASH160 $\langle h_p \rangle$ OP_EQUALVERIFY
OP_CHECKSIG

and the corresponding unlocking script would have to be:

$$\langle sig \rangle \langle P \rangle \langle d \rangle.$$

However, this restricts who will be able to redeem it to the owner of the public key P . It is recognized herein that this may not be desirable in some applications, e.g. where Alice wishes to retain the ability to designate signatory authority only after setting up the puzzle.

It is recognized herein that hash puzzle functionality can be emulated by exploiting the r-part in an ECDSA signature, which may be an ephemeral random value. The ECDSA signature

consists of two main parts, r and s . As seen above, $r = [k \cdot G]_x$. In place of a conventional hash puzzle $h = H(d)$, the intractability of inverting elliptic curve addition can form an analogous puzzle called herein an r -puzzle. To solve the puzzle, one would need to obtain the solution value k , where k is the ephemeral key corresponding to r .

5

With conventional hash puzzles, the risk is revealing d onto the blockchain when solving the puzzle. However, with the r -puzzle, k is never revealed. Instead r is revealed and from r along with the signature, the knowledge of k can be proved.

- 10 To emulate hash puzzle functionality, the creator of the r -puzzle may first hash some other pre-image data to get the value k , since k must be a fixed size whereas the pre-image data of a hash puzzle can be any length (and one property of a hash function is that it outputs a value of a fixed length regardless of the length of the input data). For example, if using private/ephemeral keys that are 256 bits long, then the pre-image data to the r -puzzle
- 15 should be hashed to get k . Alternatively however, some suitable-length value of k could just be selected and used as the secret value directly in its own right (i.e. there is no need to derive it from some other, preceding pre-image).

- This method can be used with any blockchain system that uses ECDSA signatures for
- 20 spending. By way of illustration, the following will describe an example implementation in a UTXO-based model. In the scripting language, the OP_CHECKSIG opcode requires a signature and a public key on the stack (with the public key on the top of the stack and the signature immediately below it). For the r -puzzle, the script is configured to check that the r value in the signature provided is the same one used for the r -puzzle challenge. In other
- 25 words, the script will not only check that the signature is valid on the public key (through OP_CHECKSIG), it will also make sure that the signature is created using the r value of the r -puzzle, which is to be published on the blockchain beforehand.

- Some example implementations of an r -puzzle are now discussed with reference to Figures
- 30 7 to 10. In each case the prover, e.g. Bob, has created a signature (r, s) by signing a part of Tx_2 . A signature of this form may also sometimes be referred to as " sig ". In the context of

cryptographic signatures, the signed part is also called the “message” (m). The signed part (message) m includes at least the output 203₂ of Tx_2 which will lock the resulting payment to Bob. If there is more than one output, m may comprise some or all of the outputs. m may also include other parts such as the locktime if used. However it will typically exclude the
 5 unlocking script itself (and of course must at least exclude the signature itself). The part of Tx_2 to be signed as the message m could be set by Sighash, or could be a default, or a fixed feature of the protocol.

Perhaps the simplest implementation of an r-puzzle is shown in Figure 7. The locking script
 10 in Tx_1 comprises a reference instance of the r-part, labelled here r' . In this method, the unlocking script in Tx_2 need only contain at least the s-part (s) of Bob’s signature. It may also include the public key P corresponding to the private key V which Bob used to sign m . The locking script of Tx_1 is configured so as, when run by the script engine 402 at a node
 104, to take s and P from the unlocking script of Tx_2 and perform the following operations:

15

- I) $R' = H_{sig}(m)s^{-1} \cdot G + r's^{-1} \cdot P$, and
- II) check $[R']_x = r'$,

where r' is taken from the locking script of Tx_1 , and s and m are taken from the unlocking
 20 script of Tx_2 . Bob’s public Key P may also be taken from the unlocking script Tx_2 , or it may be known by other means. H_{sig} is a hash function that was used to hash m in generating the first ECDSA signature. It may be any form of hash function. Whatever form it takes, the form (type) of this hash function may be assumed to be predetermined and known at both ends. G is a fixed, publicly known vector value.

25

The locking script is configured to return the result of “true” on condition that said check is true, but to return a result of “false” otherwise. In the UTXO case, a true (i.e. successful) outcome of running the locking together with the unlocking script is a requirement for validity of the transaction. Thus the validity of the Tx_2 can be used as a proxy for the
 30 outcome of the r-puzzle. Or put another way, the validity of Tx_2 is conditional on providing the solution to the r-puzzle. I.e. if Bob does not pass the r-puzzle, his transaction Tx_2 will

not be propagated over the network 106 nor recorded in the blockchain 150 (and any payment defined in the output of Tx_1 will not be redeemed).

Whilst the example of Figure 7 may be simplest in a mathematical sense, this does not necessarily mean it is simplest to integrate with any given node protocol or scripting language. If the spender only provides $\langle s \rangle$ and $\langle P \rangle$ in the unlocking script as opposed to $\langle r, s \rangle$ and $\langle P \rangle$, then the script must account for this. Operations I)-II) are not the operations of a standard Checksig type opcode. The OP_CHECKSIG op-code expects the signature to be in DER format so if only the $\langle s \rangle$ value is provided in the unlocking script then there will need to be some additional op-codes in the locking script (OP_CAT to concatenate etc.) in order to produce a valid signature in DER format. Figure 8, described shortly, shows an alternative example that, whilst involving an extra step mathematically speaking, in fact integrates more simply with scripting languages such as Script that already have a dedicated opcode for invoking an ECDSA signature verification based on r and s both being taken from the input of Tx_2 .

Note also: it is not essential to include P in Tx_2 in all possible embodiments. In fact, from knowledge of the message m and (r, s) , or in this case (r', s) , it is possible to compute two possible values P and $-P$ of the public key (but not to know which is which). Two verifications can then be used identify which is the correct one, or alternatively a one bit flag can be included in Tx_2 to signal which of the two possible solutions to use. This latter approach is currently used in some account-based protocols. However it tends not to be used in current UTXO-based protocols where the scripting language (e.g. Script) has no opcode for the operation to compute P and $-P$ from (r, s) and m . Nonetheless, the possibility should not be excluded that one could be introduced or that the operation could simply be explicitly coded into the locking script. Another possibility is that Alice already knows or has access to P or receives it over the side channel 301. However that would require a separate look-up to map P to Tx_2 .

Another example implementation is shown in Figure 8. Here the r-puzzle requires that the unlocking script of Tx_2 explicitly includes a submitted instance r of the r-part. The locking

script of Tx_1 comprises a test for the r-part, the test comprising a reference instance r' of the r-part to be compared against the submitted instance r . In this method, the unlocking script in Tx_2 must contain at least the r-part (r) and the s-part (s) of Bob's signature. It may also include the public key P corresponding to the private key V which Bob used to sign m .

- 5 The locking script of Tx_1 is configured so as, when run by the script engine 402 at a node 104, to take r , s and P from the unlocking script of Tx_2 and perform the following operations:

- 10 I) check $r' = r$, and
 II) compute $R' = H_{sig}(m)s^{-1} \cdot G + rs^{-1} \cdot P$, and
 III) check $[R']_x = r$,

- where r' is taken from the locking script of Tx_1 , and s , r and m are taken from the unlocking script of Tx_2 . Bob's public Key P may also be taken from the unlocking script Tx_2 , or it may
 15 be known by other means, such as by being derived from (r, s) and m or (r, s) and m as discussed previously.

- The locking script is configured to return the result of "true" on condition that the checks in both steps I) and III) are true, but to return the result of "false" otherwise. Again in the
 20 UTXO based case, this enables the validity of the transaction to be determined in dependent on the outcome of the r-puzzle knowledge proof. Note that the numerals I-III do not necessarily imply an order. The check I) could be performed before or after II)-III), though III) does have to be performed after II).

- 25 In the method of Figure 8, steps II) and III) alone are the conventional operations performed by the ECDSA verification function. In most protocols they can therefore be invoked by a dedicated opcode such as the existing Checksig opcode (OP_CHECKSIG) in Script. Step I) can be coded separately into the locking script using general purpose opcodes (an example is given shortly). It is also not excluded that steps II) and III) could in principle be explicitly
 30 encoded using general purpose opcodes instead of using a dedicated opcode such as Checksig.

In one example transaction protocol, transaction ECDSA signatures use the ASN.1 (Abstract Syntax Notation One) DER (Distinguished Encoding Rules) encoding format, as shown in Figure 12. The first byte field contains the flag 0x30 denoting the ASN.1 sequence number.

- 5 The next byte field contains the length of the sequence in hexadecimal. The third byte field contains the flag 0x02 denoting an ASN.1 integer. After that, the r value of the ECDSA signature is contained in the next 32 or 33 bytes. The field should be 32 bytes however if the first byte of r is greater than 0x7f (the first bit is a one), then an addition byte of zeros is added in front of the r value, making it 33 bytes long. This is done as a result of the DER
- 10 format encoding which interprets the first bit of the integer as the sign. The extra byte of zeros is added to the beginning of the value so that it is not interpreted as a negative value. The same thing is done of the s value of the ECDSA signature. Finally, a one-byte field, hash-type (ht), is added to the DER encoding which corresponds to the type of transaction signature in the transaction (SIGHASH_ALL, SIGHASH_NONE, etc.).

15

Consider the case where Alice (A) wants to create an r -puzzle transaction in which anyone who obtains the solution to the puzzle can spend. To achieve this, she will create a new transaction Tx_1 like the one shown below. The inputs section includes the unlocking script of the previous transaction Tx_0 being spent. For simplicity, assume it to be a standard

20 P2PKH that is spent using Alice's signature and public key. The outputs section includes the *locking script* (script pub key), or in other words the r -puzzle challenge. As shown in Figure 12, signatures may use the DER encoding format in some protocols, so the script must extract the value of r from the encoded signature and then check that it is equal to $\langle r \rangle$.

- After that, the script must check that the signature is valid on the public key. A more
- 25 detailed description of how the script works is shown in figure 5. The op-codes in **bold** are essentially just a way of extracting r from the signature.

$TxID1$	
Inputs	Outputs
<i>Any spending inputs</i>	OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT OP_DROP $\langle r \rangle$ OP_EQUALVERIFY OP_SWAP OP_CHECKSIG

The corresponding *unlocking script* is shown below, where the signature sig_r uses r and the spender Bob (B) can compute the signature using any private/public keypair. Note that sig_r is (r, s) .

$$\langle P_B \rangle \langle sig_r \rangle$$

5

Figure 13 shows a step-by-step script analysis.

The ephemeral key k may be generated by Alice and given to Bob (and optionally one or more other potential provers). Alternatively k may be generated by Bob and given to Alice
 10 to set an r -puzzle that only Bob can solve (or anyone Bob chooses to share k with). In either case, the prover Bob must trust the sender Alice not to spend the transaction herself since she knows the solution (k) to the r -puzzle. To prevent this, the prover Bob could create the puzzle and then send the r value to Alice for her to use when creating the R -puzzle transaction. After that, Bob can redeem the output at a later date using any private/public
 15 keypair, as long as he keeps the value k , which is the solution to the r -puzzle and can be seen as a form of key. On the other hand, in some cases the fact that Alice knows k can be an advantageous feature. For example this can be used to create a private key puzzle, and through that a generalised atomic swap.

20 Figure 9 shows another example of an r -puzzle, which may be termed herein a “pay to r -puzzle hash” (P2RPH), by analogy with pay to public key hash (P2PKH). For added security and privacy, the r value can be hashed before being placed in Tx_1 (which will be propagated thorough the nodes 104 of the network 106 and placed on the blockchain 150). Similar to P2PKH, where only a hash of the public key is on the blockchain instead of the public key
 25 itself, the same can be done with R -puzzles.

Here the r -puzzle again requires that the unlocking script of Tx_2 includes a submitted instance r of the r -part. The locking script of Tx_1 again comprises a test for the r -part, but this time in the form of a compressed instance of the r -part in the form of a hash of r' , i.e.
 30 $h = H(r')$. This will be compared against the submitted instance r . In this method, the unlocking script in Tx_2 must again contain at least the r -part (r) and the s -part (s) of Bob’s

signature. It may also include the public key P corresponding to the private key V which Bob used to sign m . The locking script of Tx_1 is configured so as, when run by the script engine 402 at a node 104, to take r, s and P from the unlocking script of Tx_2 and perform the following operations:

5

- I) check that $h = H_{puz}(r)$, and
- II) compute $R' = H_{sig}(m)s^{-1} \cdot G + rs^{-1} \cdot P$, and
- III) check $[R']_x = r$,

- 10 where h is taken from the locking script of Tx_1 , and s, r and m are taken from the unlocking script of Tx_2 . The hash value $h = H_{puz}(r)$ where H_{puz} is the hash function used in the hash-of-r puzzle. It may be any form of hash function. It may be the same or a different form of hash function to H_{sig} . Whatever form it takes, the form of H_{puz} may be assumed to be predetermined and known at both ends. Bob's public Key P may also be taken from the
- 15 unlocking script Tx_2 , or it may be known by other means, such as by being derived from (r, s) and m or (r, s) and m as discussed previously.

- The locking script is configured to return the result of "true" on condition that the checks in both steps I) and III) are true, but to return the result of "false" otherwise. The check I) could
- 20 be performed before or after II)-III), though III) does have to be performed after II).

- Also, again just as in the case of Figure 8, steps II) and III) alone are the conventional operations performed by the ECDSA verification function. In most protocols they can therefore be invoked by a dedicated opcode such as the existing Checksig opcode
- 25 (OP_CHECKSIG) in Script. Step I) can be coded separately into the locking script using general purpose opcodes.

An example of the locking script in the transaction challenge Tx_1 is shown below:

$TxID1$	
Inputs	Outputs

<i>Any spending inputs</i>	OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT OP_DROP OP_HASH160 $\langle h \rangle$ OP_EQUALVERIFY OP_SWAP OP_CHECKSIG
----------------------------	---

Any type of hash function could have been used which is consistent between both parties, the sender and the receiver. However, staying consistent with the P2PKH standard, we use OP_HASH160, a double hash of SHA-256 and then RIPEMD-160.

5

The corresponding *unlocking script* is shown below (the same as in the previous section), where the signature sig_r uses r and the spender Bob (B) can compute the signature using any private/public keypair:

$$\langle P_B \rangle \langle sig_r \rangle$$

10

The example of Figure 9 is thus just like Figure 8 except that it uses a hash of the r -part as the basis of the r -challenge instead of an untransformed instance of r .

Note that in any of these cases, it is not excluded that the unlocking script of Tx_1 could impose additional criteria for a “true” outcome. E.g. an example would be a locktime or a requirement for an additional signature.

An example use case of any of the above techniques is as a general knowledge challenge. Consider any challenge that has some solution k , or some solution that can be hashed to k . Alice can then create an R-puzzle that is coupled to the puzzle. Namely, she can define $r = [k \cdot G]_x$.

As an example, Alice is a maths professor. She can construct an r -puzzle transaction Tx_1 , where the underlying k value is the solution to a maths question which students are incentivised to solve. Whoever works out the solution can use that to create a signature (r, s) , where r will match the value in the locking script, hence, claim the reward. The signature not only provides authenticity but also acts a knowledge proof of the solution without revealing the solution to anyone else. R-puzzles thus provide a secure mechanism to prove the knowledge of some solution or information in general without the risk of

25

exposing it. It elegantly reuses signatures required in unlocking scripts and allows whoever finds the solution to claim the reward with privacy, as any public key P can be used.

This scheme can also be used as a form of token or digital ticket. For example, an event
 5 organiser can issue different values of k as digital tickets to the attendees. When an attendee wants to attend the event, they can prove knowledge of the secret token through the use of the r-puzzle.

As another example use case, an r-puzzle can be used as a signatory authorisation scheme,
 10 where one party can delegate the right to sign to another party. Consider an r-puzzle transaction Tx_1 that can only be unlocked if a signature with an r value that matches the locking script is provided. This implies that only a person who knows the value k , where $[k \cdot G]_x = r$ can produce a such signature. However, if the person passes on the knowledge of k to someone else, then this is effectively authorizes the other person to sign on his or
 15 her behalf.

For example, suppose Alice wants to receive a delivery. She is worried that she might not be there to accept the delivery. She gives both Bob and Charlie a copy of k so that they can accept the delivery on her behalf. If Dave is delivering the parcel, she must get a signature
 20 with the expected r value to release the parcel to Bob.

In a scenario like this, k can be thought of as acting as an ephemeral private key, and r as an ephemeral public key; analogous to V and P respectively, except that k and r are not linked to a particular identity.

25

JOINT-VALUE R-PUZZLE

As an extension to the hashed R-puzzle (P2RPH) of Figure 9, it is possible to include an extra value d concatenated with r before hashing (to get $h = H_{puz}(r||d)$). In that case, the
 30 prover (e.g. Bob) must not only solve the r-puzzle, but also know d . An example of this is shown in Figure 10.

The locking script of Tx_1 is configured so as, when run by the script engine 402 at a node 104, to take r, s, P and d from the unlocking script of Tx_2 and perform the following operations:

5

- I) check $h_{joint} = H_{puz}(r||d)$, and
- II) compute $R' = H_{sig}(m)s^{-1} \cdot G + rs^{-1} \cdot P$, and
- III) check $[R']_x = r$,

- 10 where $r||d$ represents a concatenation of r and d in either order (r first or d first). An example of the locking script in the challenge transaction Tx_1 is shown below:

$TxID$	
Inputs	Outputs
<i>Any spending inputs</i>	OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT
	OP_DROP OP_2 OP_ROLL OP_CAT OP_HASH160 $\langle h_{joint} \rangle$
	OP_EQUALVERIFY
	OP_OVER OP_CHECKSIGVERIFY OP_CHECKSIG

The corresponding unlocking script is shown below (the same as in the previous section except with d included). The signature $sig_r P_B$ uses r and the prover Bob (B) can compute the signature using any private/public keypair.

15

$$\langle sig' \rangle \langle P_B \rangle \langle d \rangle \langle sig_r \rangle$$

The extra signature sig' is an added feature for security (see the section on optional security features later). However this need not be required in all possible embodiments.

20

An example use case would be a CLTV Linked R-Puzzle. In this case, the data value d can be a time value t which is linked to a CLTV (Check Lock Time Verify) transaction output. The motivation behind this is to hide the time t that the output cannot be spent before within the P2RPH hash and link it to an R-Puzzle. In that case, the prover (e.g. Bob) must not only

25

solve the r-puzzle, but also know t and wait until the specific time to spend it. An example of the locking script in the transaction is shown below:

$TxID1$	
Inputs	Outputs
<i>Any spending inputs</i>	OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT OP_DROP OP_2 OP_ROLL OP_CHECKLOCKTIMEVERIFY OP_CAT OP_HASH160 $\langle h_{joint} \rangle$ OP_EQUALVERIFY OP_OVER OP_CHECKSIGVERIFY OP_CHECKSIG

The corresponding *unlocking script* is shown below, where the signature $sig_r P_B$ uses r and the spender Bob (B) can compute the signature using any private/public keypair.

$$\langle sig' \rangle \langle P_B \rangle \langle t \rangle \langle sig_r \rangle$$

The extra signature sig' is an added feature for security (see the section on optional security features later). However this need not be required in all possible embodiments.

The above has been described in terms of a concatenation. However, it is also possible to generalize this to some function $f(r, d)$. For example f could be an addition of r and d , e.g. implemented as $\langle r \rangle \langle d \rangle$ OP_ADD.

MULTIPLE R-VALUE STATEMENTS

Another possibility is to have multiple pre-determined values of r , say r_1, r_2 and r_3 which are associated with and unlock different statements. If one assigns a statement S_i to each r_i , then we can acknowledge a particular statement by using the corresponding r_i in the signature. For example, this may be used to sign to indicate to consent to one or multiple alternative possible terms in agreement.

It is possible to construct a locking script that checks which r value is used in an unlocking script, and one can assign interpretations to values of r . A locking script that implements the above idea may look like this:

```

OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT OP_DROP
OP_DUP OP_HASH160  $\langle H(r_1) \rangle$  OP_EQUAL
OP_IF
     $\langle \text{Statement 1} \rangle$  OP_DROP OP_OVER OP_CHECKSIGVERIFY
5  OP_ELSE
    OP_DUP OP_HASH160  $\langle H(r_2) \rangle$  OP_EQUAL
    OP_IF
         $\langle \text{Statement 2} \rangle$  OP_DROP OP_OVER OP_CHECKSIGVERIFY
    OP_ELSE
10     OP_HASH160  $\langle H(r_3) \rangle$  OP_EQUAL
        OP_IF
             $\langle \text{Statement 3} \rangle$  OP_DROP OP_OVER OP_CHECKSIGVERIFY
        OP_ENDIF
    OP_ENDIF
15 OP_ENDIF
OP_CHECKSIG

```

Every $\langle \text{statement } i \rangle$ is to be replaced by different locking conditions which can only be accessed after solving the corresponding R-Puzzle. The unlocking script is shown below with

20 r_i being the distinct r value needed to access on of the set statements.

$$\langle \text{sig}' \rangle \langle P \rangle \langle \text{sig}_{r_i} \rangle$$

The extra signature sig' is again an optional added feature for security (see later). However

25 this need not be required in all possible embodiments.

OPTIONAL SECURITY FEATURE #1

If a signature based on k gets published, then anyone who knows the value of k can derive

30 the value of the secret key V used to create the signature. This can be done by solving for V in the signature equation below.

$$s = k^{-1}(H(m) + rV) \bmod n$$

Solving for V we get:

$$V = r^{-1}(sk - H(m)) \bmod n$$

- 5 This does not pose significant risks as in many cases the receiver of the transaction is the only one who knows k . In other cases, the spender must be wary never to re-use the private key V which was used to sign the solution to the R-puzzle. Good security practice dictates that it is preferable for a user never re-use public/private keypairs (P, V) , but rather always use a fresh new public/private keypair when receiving new money.

10

In principle, the public-private key pair (P, V) is “permanent”. That is, it can be used many times. The use of the random ephemeral key k should ensure this. However, there have been incidents where the random number generator was poorly implemented.

- If one signs two different messages using the same ephemeral key k and the same private key, then one can derive the private key V from the two signatures. I.e. given (r, s) and k ,
 15 one can work out V , where $r = [k \cdot G]_x$ and V is the private key to the public key P used in the signature. If the random number generator fails during the signing process, it may generate the same random number as last time, hence leaking the private key to the public. To address the problem, people start to avoid reusing public keys instead of fixing random
 20 number generators.

- In the present case, if Alice knows k , but she does not know V , the private key to Bob’s public key. When Alice passes on k to Bob. Bob will be able to solve the r-puzzle by providing (r, s) using his private key. When Alice sees the signature, as she knows k , she will
 25 be able to derive V . This might not be desirable for Bob. Therefore Bob should preferably avoid re-using (P, V) .

- However, an issue with this is that Bob’s public key P can then not be used as a persistent means of identifying Bob.

30

To address this, according to embodiments disclosed herein, Bob may include an additional signature sig_2 of Bob in Tx_2 using a separate private key V_2 having a corresponding public key P_2 . He also includes P_2 in along with the extra signature. There are thus two types of public-private key pairs. The first type is those which are generated on the fly for one-time use. The other type is those which are generated according to some extra protocols, e.g., HD wallet. Bob can use the first type of key pair for r puzzle signature, and use the second type for the second signature.

Alice can then use the second public key to look up an identity of Bob, e.g. a proper name, username, or network address of Bob, based on a mapping between the public key and the identity. The mapping could for example be made available in a public database mapping public keys to identities, or the mapping could simply be pre-agreed between Alice and Bob (e.g. stored privately on Alice's computer equipment 102a).

Consider again the signatory authority use case. For example, Alice wants to receive a delivery but might not be available accept the delivery herself. She gives both Bob and Charlie a copy of k so that they can accept the delivery on her behalf. Dave is delivering the parcel. He must get a signature with the expected r value. Now imagine that for his records or regulatory compliance, Dave also needs to verify the identity of the receiver.

Suppose Bob is there to accept the delivery. If Bob produces his public key and a signature based on k , then both Alice and Charlie will be able to work out Bob's private key V . This is not an issue if the public key is designed for one time use only. However, it will not be ideal if Bob needs this public key to prove his identity in the future.

To address this issue, embodiments may include in Tx_2 one more signature that is independent of the r-puzzle from Bob which can be used to identify Bob. For example, the extra signature and the corresponding public key P_2 can be added to an OP_RETURN output (an unspendable output) in the same transaction that Dave accepts. An alternative is to include an extra OP_CHECKSIG in the locking script of the r-puzzle transaction. By browsing

the transaction and the public key used for the extra signature, Alice can tell who has signed on her behalf.

- In some other cases, there can be concerns that the value k might be leaked prior to use. To address this, Alice can add a P2PKH to an r-puzzle transaction to make it securer. Suppose Alice would like to delegate her signing right to Bob. Alice obtains a one-time public key P_2 from Bob and creates an r-puzzle transaction that not only specifies the r value but also specifies the extra public key P_2 .
- 10 In order for Alice herself to be able to sign as well, optionally Alice can create a 1-out-of-2 MultiSig. An example of the locking script is given below:

<i>TxID</i>	
Inputs	Outputs
<i>Any spending inputs</i>	<code>OP_DUP OP_3 OP_SPLIT OP_NIP OP_1 OP_SPLIT OP_SWAP OP_SPLIT</code> <code>OP_DROP $\langle r \rangle$ OP_EQUALVERIFY</code> <code>OP_OVER OP_CHECKSIGVERIFY OP_CHECKSIGVERIFY</code> <code>OP_1 $\langle \text{Alice's PK} \rangle \langle \text{Bob's } P_2 \rangle$ OP_2 OP_CHECKMULTISIG</code>

- Note that the r-Puzzle provides more flexibility as Alice can choose when to pass the solution of the r-puzzle, i.e., the signing right, to Bob. She can decide to pass on or not to pass on even after the transaction is mined.
- 15

- If k is leaked, then people can discover the private key that is used to sign the signature with the leaked k . However there is another private key V_2 : the private key that is linked to the public key that can be used to identify Bob. For the output to be compromised, the attacker has to obtain two independent secrets which is much more unlikely than compromising only one of them.
- 20

- Note, in the above example, the locking script of Tx_2 is locked to Bob's extra public key P_2 by means of a conventional P2PKH (to be unlocked by the extra signature, not the one used in the r-puzzle). The r-puzzle technique allows an additional choice for users. In some applications it may be desired to use the r-puzzle so that prover is allowed to meet the challenge, regardless of identity. In some other applications on the other hand, the
- 25

combination of hash puzzle and P2PKH may still be desirable, and the r-puzzle can optionally be used in conjunction with that. This will be discussed in more detail later.

If however the extra signature corresponding to P_2 is required for identity look-up and/or security, but *without* the locking script of Tx_1 being tied in advance to the identity of a particular prover as in a P2PKH, then the above locking script can be adapted accordingly. That is, it can simply include a Checksig on the extra signature, but *not* an OP_EQUALVERIFY on the corresponding public key P_2 .

10 HASH COLLISION BOUNTIES

Hash collision bounties could be implemented on the blockchain by constructing a transaction output of the kind shown in Table 1 which can be redeemed (spent) by anyone who can find a collision to a hash function (SHA1 in the example below).

<i>TxID</i>	
Inputs	Outputs
<i>Any spending inputs</i>	OP_2DUP OP_EQUAL OP_NOT OP_VERIFY OP_SHA1 OP_SWAP OP_SHA1 OP_EQUAL

Table 1: Hash Collision Bounty Transaction

The bounty can be “claimed” (i.e. the output of the hash bounty transaction may be spent), by a proof transaction with the corresponding *unlocking script*:

$$\langle preimage_1 \rangle \langle preimage_2 \rangle$$

The bounty rewards whoever finds two different preimages that are hashed to the same value.

The problem with the bounty is that whoever is trying to claim the reward has to reveal the solution. Therefore, anyone who sees the proof transaction will be able to hijack the answer and create their own proof transaction to claim the bounty for themselves. Moreover, a malicious attacker would be free to replicate the attack, which may pose a security risk.

This issue is resolved using a form of r-puzzle transaction for preimage bounties, which constructs a zero-knowledge proof on the r-part of an ECDSA signature.

A hash collision bounty transaction effectively rewards anyone who can derive a triplet
5 (x_1, x_2, h) such that $x_1 \neq x_2$ and $H(x_1) = H(x_2) = h$, where H is the hash function in the bounty.

The party constructing the bounty transaction has no knowledge of what values (x_1, x_2, h) may take, so these cannot be specified in the bounty transaction.

10

The solution uses a zero-knowledge proof hash circuit, applying the mythology set out in International Application No. PCT/IB2019/052184 [1], which is incorporated herein by reference in its entirety.

15 Further details of example proof verification algorithms $\mathcal{V}$ are described below.

Figure 14 shows a bounty transaction 1402 (the challenge transaction in this example) and a corresponding proof transaction 1404. The bounty transaction 1402 rewards proof of a successful hash collision attack as set out below.

20

In the following examples, the hash function is defined to be the SHA256 to be the hash function. This is defined in the locking script of the bounty transaction 1402.

By applying the methodology of [1], a zero-knowledge proof is constructed on the following
25 statement:

“the preimage of h is the same as the ECDSA private key to the ECDSA public key R ”.

The proof is embodied as zero-knowledge proof component π included in the proof
30 transaction, which proves the above statement in respect of the r-part of the ECDSA signature and the hash value h (also included in the proof transaction).

A proof verification algorithm $\mathcal{V}$ takes three inputs (h, R, π) and outputs 1 (true) if π is a valid proof on (h, R) , and false otherwise. When $\mathcal{V}$ outputs 1, it means that the prover knows a secret value d such that $H(x) = h$ and $d \cdot G = R$. The proof verification algorithm

5 $\mathcal{V}$ is denoted in pseudo-code as `[ZK_VERIFY]`. This algorithm encodes one or more proof requirements which must be satisfied by the zero-knowledge proof component π of the proof transaction with respect to R and π .

Applying the terminology associated with EDCSA, it can be seen that d fulfils the role of

10 the ephemeral key (i.e. $d = k$ in the notation introduced above).

Assuming there exists a SHA256 collision (d_1, d_2, h) where all three values are unknown, let

$$R_1 = d_1 \cdot G$$

$$R_2 = d_2 \cdot G$$

15 Note that $d_1 \neq d_2$ implies that $R_1 \neq R_2$.

A prover must to provide two proofs (π_1, π_2) , where π_1 is a proof on (h, R_1) and π_2 is a proof on (h, R_2) – note both proofs are for the same hash value h which, in the present context, is not known in advance and is not specified in the bounty transaction or anywhere

20 else in the blockchain (or otherwise).

The bounty transaction 1402 is an R-puzzle transaction for a SHA256 collision bounty constructed with a locking script which may be represented in pseudo-code as follows:

```

25 [EXTRACT_r] [ZK_VERIFY]
   OP_IF
     [EXTRACT_r] [ZK_VERIFY]
     OP_IF
       [EXTRACT_r] [EXTRACT_r] OP_EQUAL OP_NOT
30       OP_IF
         OP_CHECKSIGVERIFY OP_CHECKSIG
       OP_ENDIF
     OP_ENDIF
   OP_ENDIF
35 OP_ENDIF

```

Here, `[EXTRACT_r]` represents a function which extracts r from an ECDSA signature (r, s) .

In the present example, the hash function H is defined in the locking script of the challenge transaction 1402 as SHA256, as part of the proof verification algorithm $\mathcal{V}([\text{ZK_VERIFY}])$.

- 5 The unlocking script in the proof transaction 1404 is:

$$\langle r_1, s_1 \rangle \langle r_2, s_2 \rangle \langle P_1 \rangle \langle P_2 \rangle \langle h \rangle \langle \pi_1 \rangle \langle \pi_2 \rangle$$

where $r_1 = [d_1 \cdot G]_x$ and $r_2 = [d_2 \cdot G]_x$.

10

It can be seen that when the locking script in the bounty transaction 1402 is applied to the unlocking script in the proof transaction 1404, a check is performed that $r_1 \neq r_2$. Assuming that to be the case, a first instance of the verification algorithm extracts r_2 , derives R_2 from r_2 , and then verify the proof π_2 on (h, R_2) . If the proof is valid, then r_1 is extracted, R_1 is
 15 derived from r_1 , and π_1 is verified on (h, R_1) .

Certain r_i values will correspond to two possible elliptic curve points, denoted R_i^+ and R_i^- .

In this case, a flag may be included in association with r_i to indicate which of the two points R_i^+ , R_i^- is intended. Alternatively, the flag may be omitted, and the verification algorithm

- 20 may be applied for both R_i^+ and R_i^- , and in that event it is sufficient for the proof to be correct in respect of one of R_i^+ and R_i^- . With two r-parts, r_1, r_2 , there may be up to four unique R -pairings, i.e. (R_1^+, R_2^+) , (R_1^+, R_2^-) , (R_1^-, R_2^+) , (R_1^-, R_2^-) - and it is sufficient for one of these to be correct, i.e. to be proved to correspond to colliding preimages.

- 25 Alternatively, the points on the elliptic curve R_1 and R_2 may be included in the proof transaction 1404 explicitly. In that event, the "r-part" of the relevant signature (r_i, s_i) is the x -coordinate of the point R_i which is used together with the y -coordinate of that point to check the proof π_i .

- 30 If the proof is also valid, then all signatures will be verified. Note that for simplicity, the steps to verify the security signatures are omitted from the locking script. In practice, they

will be included in the locking script in a manner that will be apparent to the skilled person. These steps verify each of the signatures (r_1, s_1) , (r_2, s_2) based on the respective public keys P_1, P_2 associated therewith (included in, or otherwise derivable from, the unlocking script as shown above), and respective signed parts of the transaction denoted m_1, m_2 . In general,

5 the signed parts may be the same (i.e. $m_1 = m_2$) or different ($m_1 \neq m_2$); likewise, the public keys P_1 and P_2 may be the same ($P_1 = P_2$) or different ($P_1 \neq P_2$) – however, particular benefits are attained when $P_1 = P_2$ and $m_1 \neq m_2$ (see below).

There is no restriction on the identity of the prover, i.e. the public keys P_1 and P_2 of the

10 proof transaction 1404 are not specified in the bounty transaction 1402 or anywhere else in the blockchain (or otherwise). That is, the bounty can be claimed by any party who provides a valid pair of signatures with different r-parts, and zero-knowledge proof components which are correct for the respective r-parts of those signatures, irrespective of whose public key or keys are associated with those signatures (i.e. irrespective of the private key or keys

15 P_1, P_2 used to verify them).

OPTIONAL SECURITY FEATURE #2 - SIGNATURE FORGABILITY

There is a potential source of vulnerability in the above method which could be exploited by

20 a miner trying to claim the funds. A miner who receives the proof transaction (from a prover) can change the transaction to send the funds to himself while using the same signature that the spender used in the original transaction. This is done as follows:

Let $P = V \cdot G$ be the public/private keypair used to sign the original transaction denoted by m to get a signature (r, s) such that:

25 $r = [k \cdot G]_x$, and
 $s = k^{-1}(H(m) + rV) \bmod n$.

To spend that transaction, the spender will use the following unlocking script:

$$\langle P \rangle \langle r, s \rangle$$

The miner who receives this transaction can change the transaction into a new one denoted by m' which sends the funds to himself using the following new unlocking script:

$$\langle P' \rangle \langle r, s \rangle$$

- 5 where $P' = V' \cdot G$ is the public/private keypair such that:

$$V' = V + r^{-1}[H(m) - H(m')], \text{ and}$$

$$P' = P + r^{-1}[H(m) - H(m')] \cdot G.$$

Note that the miner does not need to know V' (since they do not know V).

- 10 The verification process is done using the following calculation:

$$R' = H(m)s^{-1} \cdot G + rs^{-1} \cdot P$$

The signature is valid if and only if $(R')_x = r$, otherwise it is invalid.

With the new transaction m' and the new unlocking script, the verification process is as

- 15 follows:

$$R' = H(m')s^{-1} \cdot G + rs^{-1} \cdot P'$$

$$= H(m')s^{-1} \cdot G + rs^{-1} \cdot \{P + r^{-1}[H(m) - H(m')] \cdot G\}$$

$$= rs^{-1} \cdot P + H(m)s^{-1} \cdot G$$

- 20 $= r$

One solution to this vulnerability is to include another extra signature sig' (anti-forgery signature) in the unlocking script on another message $m_{sighash}$ or the same message m which a miner will not be able to provide unless they know the secret key V . In that case

- 25 the unlocking script would include:

$$\langle sig' \rangle \langle P \rangle \langle sig_r \rangle$$

sig' must be a signature on a different message m' e.g. $m' = m_{sighash}$ but generated using the same public key V . In a UTXO blockchain context a different “sighash flag” may be used to achieve this (for example `SIGHASH_NONE` instead of `SIGHASH_ALL` which is the default flag).

- 30

Also, sig' must use a different value of r (i.e. a different ephemeral key) so that it does not leak the private key (since the private key can be derived from two signatures on different messages but which use the same ephemeral key and the same public key - a known property of ECDSA).

5

The unlocking script would, in this case, be modified to include an additional check of the anti-forgery signature sig' .

For a proof transaction which claims a hash collision bounty, this can be implemented with three signatures – two which prove the collision attack, and a third, separate anti-forgery signature.

However, to avoid the use of a third signature, the two signatures (r_1, s_1) and (r_2, s_2) in the proof transaction 1404 can be configured to simultaneously prove the collision attack and protect against forgery. This is achieved by generating both signatures using the same private key (with the consequence that the associated public keys are the same, $P_1 = P_2$), and the two signatures may or may not be generated on different parts of the transaction (i.e. with $m_1 \neq m_2$ or with $m_1 = m_2$). Note that it is already a requirement for these signatures to use different ephemeral keys in order to claim the bounty. This is an equally robust solution but with two signatures instead of three.

In that case, the public key P_1 is still not specified in advance (meaning anyone with the solution can claim the bounty), but the unlocking script does require the signatures to be valid for the same public key (whoever that may belong to).

25

HASH PREIMAGE BOUNTIES

A challenge transaction can also be provided which specifies the hash value h and can be redeemed with a proof transaction that proves knowledge of a single pre-image. This may be referred to as a hash preimage bounty transaction.

30

For example, the challenger may create a hash preimage bounty transaction for a hash value h , the preimage d of which is secret to the challenger (i.e. which the challenger has not revealed to anyone else). In those circumstances, the fact that someone else has managed to derive a preimage to h independently (without being told the preimage d by the
5 challenger) constitutes a successful preimage attack.

USE CASES

As noted, once the existence of an attack is known, that knowledge can be used to identify
10 and mitigate vulnerabilities in a cryptographic system that relies on the hash function in question.

For example, a digital certificate which depends on the security of the hash function may be revoked.
15

In a blockchain context, a transaction with one or more unspent outputs based on the hash function in question (such as a “hash puzzle” transaction, which requires the preimage to a specified hash function to be provided in order to redeem those outputs) could be redeemed by, say, the party who published the transaction once the vulnerability in the
20 hash function has been exposed, in order to prevent a malicious party from exploiting the vulnerability to illegitimately redeem those outputs for themselves.

PROOF VERIFICATION ALGORITHM $\mathcal{V}$

25 The notation π is used in the following description to denote a zero-knowledge proof component of the kind described. All description pertaining to π applies equally to π_1 and π_2 . Likewise, the notation k is used to denote a secret value which, in the context of r-puzzle, in as ephemeral key, and all description pertaining to k applies equally to k_1 and k_2 above. Likewise, a transaction signature is referred to and denoted by (r, s) , and all
30 relevant description applies equally to (r_1, s_1) and (r_2, s_2) . All description pertaining to R applies equally to R_1 and R_2 .

One aspect which must be proved is that the prover knows a value k which is a preimage of the hash value h with respect to the hash function H . This is proved in zero knowledge in that – whilst the hash function H and hash value h are known (the latter being provided in the proof transaction 1404 in the case of hash collision bounties) – k itself (i.e. the preimage) remains secret. According to the terminology used in the art, k may be referred to in this context as a “witness” (w). The statement “ k (the witness w) is a preimage of h with respect to H ” is said to be proved in zero knowledge if it can be proved without revealing the witness w . Note that, in the present context, this is only a first of two aspects which must be proved in zero-knowledge – the second aspect, which must also be proved in zero-knowledge, is that k is also the ephemeral key which has been used to generate the ECDSA signature (r, s) .

There are a number of existing mechanisms which can be applied to prove the first aspect in zero-knowledge, with respect to various hash functions H (including SHA256). Some such mechanisms make use of “arithmetic circuits”.

For example, zero knowledge “succinct non-interactive arguments of knowledge” (zkSNARKs). provide a method of proving, in zero-knowledge, the validity of arbitrary computations that can be expressed as arithmetic circuits. Two properties of zkSNARKs are that they are non-interactive (the prover sends the proof to the verifier in one move) and succinct (the proof is small and easy to verify). An arithmetic circuit is a logical circuit formed of “gates” and “wires” connecting the gates. The hash function H in question is implemented as an arithmetic circuit. Such arithmetic circuit implementations of hash functions are known in the art. Various libraries exist for implementing SNARKs and will be familiar to those skilled in the art.

By way of further example “Zcash” is a known mechanism which applies zk-SNARKs to hash functions, in order to provide knowledge of a preimage of a given hash value h without revealing h . This is based on an arithmetic circuit implementation of the hash function H in

question. For example, a known implementation of SHA256 uses 27,904 arithmetic gates, and has been applied in the context of Zcash.

The second aspect, namely equivalence of the preimage and the ephemeral key, could be proved within this framework by implementing both the hash function H and the elliptic curve scalar multiplication operation " $\cdot$ " (equivalently denoted $\times$ below) as arithmetic circuits, and applying a zero-knowledge proof protocol (such as a Sigma protocol – see below) thereto. Given an arithmetic circuit implementation of " $\cdot$ ", it will be apparent to those skilled in the art in view of the teaching presented herein, how this may be used to construct such a proof.

However, to avoid the need for an arithmetic circuit implementation of elliptic curve scalar multiplication, the approach set out in International Application No. PCT/IB2019/052184 [1], referred to above, may be applied instead. This provides a methodology which can be applied to efficiently provide both required aspects, i.e. in addition to proving knowledge of the preimage of h , proving that said preimage corresponds to the ephemeral key used to generate the signature (without requiring an arithmetic circuit implantation of elliptic curve scalar multiplication).

The aspects of that methodology which are relevant in the present context, and the background to that methodology, is set out below.

Σ -Protocols

Σ (Sigma) protocols are a type of interactive zero-knowledge proof system, that involve a number of *moves* (communications) between the prover and verifier. Usually Σ protocols involve 3 moves: the prover sends an initial commitment to the verifier (a), the verifier then responds with a random challenge (ξ) and finally the prover answers with a final response, or 'opening' (z). The verifier then accepts or rejects the statement based on the transcript (a, ξ, z) .

Σ protocols can be used to prove knowledge of, or statements about, a *witness* (w) that is known only to the prover. The protocol is zero-knowledge if it does not reveal any information about the witness to the verifier, except for the fact that a statement related to the witness is true. For further details, see Bootle, Jonathan, et al. "Efficient zero-

5 knowledge proof systems." *Foundations of Security Analysis and Design VIII*. Springer, Cham, 2015. 1-31, [Bootle 2015], which is incorporated herein by reference in its entirety.

Pedersen Commitments

10 Commitment schemes are a central part of many cryptographic protocols, and are a component of interactive zero-knowledge protocols for circuit satisfiability. A *commitment* enables a prover to *commit* to a secret value in advance, and then later verifiably reveal (*open*) the secret value. A commitment scheme has two main properties: 1) it is hiding: that is the commitment keeps the value secret, and 2) it is binding: the commitment can only be

15 opened to the originally committed value.

An example of a scheme which may be employed in the present context is the Pedersen Commitment [Bootle 2015]. This scheme involves two elliptic curve generator points: G and F in the group $\mathbb{G}$ of prime order p , known to all parties. The committer generates a secure

20 random number ρ in the field $\mathbb{F}_p$, and then computes the commitment (via elliptic curve addition/multiplication) to a secret value σ :¹

$$Com(s, r) = \sigma \times G + \rho \times F$$

The committer can at a later stage *fully open* the commitment (i.e. it can be verified), by providing the values ρ and σ . The committer can also open the commitment in response to

25 a specific challenge value as part of a Sigma protocol (without revealing ρ or σ).

A useful property of Pedersen commitments is that they are additively homomorphic, meaning that adding (on the elliptic curve) two commitments results in a commitment to the sum of the committed values, i.e.:

30
$$(\sigma_1 \times G + \rho_1 \times F) + (\sigma_2 \times G + \rho_2 \times F) = (\sigma_1 + \sigma_2) \times G + (\rho_1 + \rho_2) \times F$$

¹ Here, $\times$ denotes elliptic curve point multiplication.

This homomorphic property is exploited in zero-knowledge proofs of arithmetic circuit satisfiability.

Proofs of arithmetic circuit satisfiability in zero knowledge

5

An arithmetic circuit (over a field $\mathbb{F}_p$) is a logical construction of arithmetic *gates* that are connected by *wires* (forming a directed acyclic graph), that is capable of performing an arbitrarily complex computation². Each gate has two input wires and one output wire and performs either a multiplication ($\times$) or addition ($+$) operation on the inputs. A complete

10 circuit has free input wires and free output wires that define the external (circuit) input and output values.

A *legal assignment* of the values of the wires is defined as a set of assigned wire values which *satisfy* the circuit, i.e. each wire is assigned a value where the output of each gate

15 correctly corresponds to the product or sum of the inputs (i.e. the gate is *consistent*).

Figure 15 shows: (a) a schematic of a multiplication gate with left (w_L) and right (w_R) wire inputs and one wire output (w_O); (b) A simple arithmetic circuit with three gates, three input wires (w_1, w_2, w_3), one output wire (w_6) and two internal wires (w_5, w_6).

20

For a given arithmetic circuit, a prover can *prove* to a verifier that they know a legal assignment for the circuit without revealing the wire values, by first committing to each wire value in the legal assignment (with Pedersen commitments) and then performing special Sigma protocols with the verifier for each gate in the circuit (which can be performed in

25 parallel), with the wire values as the witness. These Sigma protocols exploit the homomorphic properties of Pedersen commitments, as described below.

To produce the proof (that a circuit is satisfied), initially the prover generates a commitment to each wire w_i in the circuit ($i = 1, \dots, m$ where m is the number of wires) and sends these

30 to the verifier:

² The computation is limited to integer operations and must have no data dependant loops or mutable state.

$$W_i = Com(w_i, \rho)$$

Σ_{zero} protocol:

- 5 For each addition gate in the circuit, the Σ_{zero} protocol is executed: this involves proving (in zero knowledge) that $w_L + w_R - w_O = 0$ (i.e. that the addition gate is satisfied: the input wires w_L and w_R equal the output wire w_O).

Note: the subscript R is a wire index and does **not** denote the point $R = k \cdot G$ in this
10 context.

1. The prover generates a commitment to zero: $B = Com(0, \rho_B)$, and sends to the verifier.
2. The verifier responds with a random challenge value: $\xi \leftarrow \mathbb{F}_p$
- 15 3. The prover then computes an opening value: $z = \xi(\rho_L + \rho_R - \rho) + \rho_B$ and sends it to the verifier.
4. The verifier checks that $Com(0, z) = \xi \times (W_L + W_R - W_O) + B$ as proof that $w_L + w_R - w_O = 0$

- 20 According to 4, in order to verify the proof, the proof verification algorithm $\mathcal{V}$ requires:
 $\mathcal{V}$ the following:

$$(z, \xi, W_L, W_R, W_O, B)$$

Note, this is for a single additive gate.

25 Σ_{prod} protocol:

For each multiplication gate the Σ_{prod} protocol is executed: this involves proving (in zero-knowledge) for each multiplication gate that $w_L \cdot w_R = w_O$ (i.e. that the multiplication gate is satisfied).

- 30 1. The prover generates 5 random blinding values: $t_1, t_2, t_3, t_4, t_5 \leftarrow \mathbb{F}_p$

2. The prover computes $C_1 = \text{Com}(t_1, t_3)$, $C_2 = \text{Com}(t_2, t_5)$ and $C_3 = t_1 \times W_R + t_4 \times F$ and then sends them to the verifier.
3. The verifier responds with a random challenge value: $\xi \leftarrow \mathbb{F}_p$
4. The prover computes the opening values:

$$\begin{aligned}
 e_1 &= w_L \xi + t_1 \\
 e_2 &= w_R \xi + t_2 \\
 z_1 &= \rho_L \xi + t_3 \\
 z_2 &= \rho_R \xi + t_5 \\
 z_3 &= (\rho_O - w_L \rho_R) \xi + t_4
 \end{aligned}$$

and sends them to the verifier.

5. The verifier then checks the following equalities:

$$\begin{aligned}
 \text{Com}(e_1, z_1) &= \xi \times W_L + C_1 \\
 \text{Com}(e_2, z_2) &= \xi \times W_R + C_2 \\
 e_1 \times W_R + z_3 \times F &= \xi \times W_O + C_3
 \end{aligned}$$

as proof that $w_L \cdot w_R = w_O$.

According to 5., the verification algorithm $\mathcal{V}$ requires:

$$(e_1, e_2, z_1, z_2, z_3, \xi, W_L, W_R, W_O, C_1, C_2, C_3)$$

Circuit proofs

The Σ_{zero} and Σ_{prod} protocols can be operated in parallel for the verification of each gate in the circuit, and the same verifier challenge value (ξ) can be used for all gates.

- 25 As an example, consider the circuit in figure 15(b): for a prover to prove in zero-knowledge to a verifier that they know a legal assignment (i.e. the wire values satisfying the circuit), the prover initially sends the wire commitments ($W_1, \dots, W_6$) and the Σ protocol commitments for each gate to the verifier (this is one additional commitment for each addition gate and five for each multiplication gate).

30

The verifier then responds with the random challenge $\xi \leftarrow \mathbb{F}_p$, and the prover computes the opening values for each gate (one for each addition and five for each multiplication) and

sends them back to the verifier. The verifier then performs the Σ protocol checks to verify that:

$$w_1 \cdot w_2 = w_4$$

$$w_4 \cdot w_5 = w_6$$

5

$$w_2 + w_3 = w_5$$

and therefore that the commitments $W_1, \dots, W_6$ correspond to *satisfying* wire values $w_1, \dots, w_6$.

10 If the prover wants to show that, in addition to satisfying the circuit, a particular wire has a particular value, they can fully open the commitments to the relevant wires. In the example, the prover can additionally send the verifier the values w_6 and ρ_6 (the verifier can then confirm that $W_6 = \text{Com}(w_6, \rho_6)$) to demonstrate that w_6 is the actual output from a particular legal assignment.

15 A simplified circuit is shown in Figure 15 (b) for the purposes of illustration. In practice useful circuits consist of many more gates. Of particular interest is an arithmetic circuit for the SHA-256 hash function - this circuit enables a prover to demonstrate that they know the pre-image (input) to a SHA-256 function that hashes to a particular (output) value, without revealing the pre-image. One of the most efficient implementations of a circuit for the SHA-256 algorithm consists of 27,904 arithmetic gates [implemented in Zcash, for example]. To
20 prove knowledge of a SHA-256 pre-image would then require the sending of ~5 MB of data in both the initial commitment and opening rounds of the above protocol, and require ~200,000 elliptic curve operations for both the prover and verifier (taking a few seconds of processor time each. There are some recently published protocols for substantially reducing
25 these computational costs and proof sizes, without introducing any restrictions on the nature of the commitments.

Efficient zero-knowledge arguments for arithmetic circuit satisfiability without pairings

30 There are several methods that have been developed to significantly improve the performance of the parallel Σ protocol approach to proving arithmetic circuit satisfiability

described in the preceding section. Reference is made to the following, each of which is incorporated herein by reference in its entirety:

Groth, Jens. "Linear Algebra with Sub-linear Zero-Knowledge Arguments." CRYPTO. Vol. 5677. 2009. [Groth 2009];

Bootle, Jonathan, et al. "Efficient zero-knowledge arguments for arithmetic circuits in the discrete log setting." Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 2016. [Bootle 2016].

The approaches described in [Bootle 2016] and [Groth 2009] involve *batching* the commitments to circuit wire values to substantially reduce the size of data that must be sent from the prover to the verifier (i.e. reducing the communication complexity). These methods enable proof systems where the communication complexity is reduced from $\mathcal{O}(n)$ to $\mathcal{O}(\sqrt{n})$ or $\mathcal{O}(\log(n))$.

Again, as a comparison for proving the satisfiability of the same SHA circuit, the protocol of [Bootle 2016] has a proving key size of just 5 KB and a key generation time of 180ms. The Proof size is 24KB and takes ~4 s to generate, and the proof also takes ~4 s to verify.

These methods in full here, except to state that the main vector batching protocol employed which is described below. This follows the same properties as the standard Pedersen commitment, but committing to n elements ($\mathbf{m} = m_1, \dots, m_n$) only requires the sending of a single group element:

1. The prover and verifier agree on a group element $F \leftarrow \mathbb{G}$
2. The prover generates n random numbers $\xi_1, \dots, \xi_n \leftarrow \mathbb{F}_p$
3. The prover computes the points $K_i = \xi_i \times F$ (for $i = 1, \dots, n$). These values form a proving key PrK that is sent to the verifier.
4. The prover generates a random value: $\rho \leftarrow \mathbb{F}_p$
5. The prover computes the commitment:

$$Com(\mathbf{m}) = \rho \times F + \sum_{i=1}^n m_i \times K_i$$

and sends it to the verifier.

Proof Verification Algorithm ($\mathcal{V}$) – Examples

5

This section describes different examples of the proof verification algorithm $\mathcal{V}$ ([ZK_VERIFY], above), and its implementation for both batched and un-batched commitment based zero-knowledge proof systems.

- 10 Two parties are involved in the zero-knowledge proof protocol: the prover (P) and the verifier (V). The purpose of the protocol is for the prover to convince the verifier that a given statement ($\mathcal{S}$) is true, while keeping information about the witness to the statement secret. The statement consists of an arithmetic circuit ($\mathcal{C}$) with q gates and m wires, as well as dependent assertions about the elliptic curve public key(s) corresponding to one (or
- 15 more) of the circuit wire values: pk_l (where the sub-script l is the wire index of the key statement. In addition, the statement may also include assertions about fully opened (public) wire values (i.e. public inputs/outputs of the circuit).

- The elliptic curve public key(s) specified in the statement correspond to a target elliptic
- 20 curve specification (which is defined by the full set of elliptic curve parameters: $\mathcal{T} = (p, a, b, G, n, h)$). In the case of Script, these parameters are defined by the specification of secp256k1. This specification includes the base generator point G .

- In addition to the specifying the base point, the statement must also specify a second point
- 25 F (where $F = f \times G$ and f is an element of $\mathbb{F}_p$). The value of f must be provably random (e.g. a hash of the genesis block 153), or a ‘nothing up my sleeve’ number, such as the first 256 bits of the binary representation of π^3 .

³ Allowing the prover a free choice over f could enable them to generate fake proofs.

Figure 16 shows an example circuit with four gates and five wires. One input wire (w_1) has its public key revealed ('opened') from the wire commitment (W_1) with the 'key-opening' value ko_1 .

5 Example 1: Individual wire commitments

This section describes *key openings* when a verifier has individual commitments to each wire in the circuit (i.e. following the Σ protocols for arithmetic circuit satisfiability – see above).

- 10 1. Each wire i ($i = 1, \dots, n$) of the circuit is committed to with a Pedersen commitment:

$$W_i = Com(w_i, \rho_i)$$

Where:

$$Com(w, r) = w \times G + \rho \times F$$

- 15 2. For the circuit wire l that requires a proof of a corresponding public key (a *key-statement proof*), the prover also sends a *key-opening*:

$$ko_l = \rho_l \times F$$

3. If a circuit wire j requires being publically revealed (a fully public wire), the prover sends a "full opening" tuple:

$$(w_j, \rho_j)$$

- 20 4. Each gate of the circuit is then proven satisfied in zero knowledge using the above Sigma protocols⁴.

5. Once the verifier has confirmed that the circuit is satisfied, the verifier then calculates the public key for the wire l via elliptic curve point subtraction:

25
$$pk_l = Com(w_l, \rho_l) - ko_l$$

6. The verifier then confirms that each pk_l matches the key(s) specified in the statement (and that the fully opened wires match specified values) to complete the validation.

⁴ This involves the prover computing and sending the Σ_{zero} and Σ_{prod} commitments (i.e. B or C_1, C_2, C_3 respectively) for each gate, the verifier replying with a challenge value (ξ), the prover sending the opening values (z and e values) and the verifier checking against the commitments.

Figure 17 shows an example protocol flow between prover (P – who provides the proof transaction 1404) and verifier (V) (challenger who sets the challenge transaction 1402) for the proof of statement $\mathcal{S}$. The statement includes the circuit description and that wire l has public key pk_l .

Example 2: Batched vector commitments

In the case of the compressed proof systems for circuit satisfiability that involve the batching of vector commitments [Bootle 2016, Groth 2009], the method described below can be used to extract *key-statement proofs* from *batched* circuit wire commitments. The full proof protocol is not described, only the generation of the batched wire commitment and the protocol to demonstrate it contains a specified public key.

15

A batched commitment is generated as follows, where the wire l is to be supplied with a key opening. m wires are batched together in the vector commitment.

1. The prover generates $m - 1$ random numbers $\xi_1, \dots, \xi_{m-1} \leftarrow \mathbb{F}_p$
2. The prover computes the elliptic curve points $K_i = \xi_i \times G$ (for $i = 1, \dots, m - 1$).
These values plus $K_m = G$ form a proving key PrK that is sent to the verifier.
3. The prover generates a random value: $\rho \leftarrow \mathbb{F}_p$
4. The prover computes the commitment to the vector $\mathbf{w}$ of wire values w_i (for $i = 1, \dots, m$) where w_m is to be key-opened:

20

$$Com(\mathbf{w}) = \rho \times F + \sum_{i=1}^{m-1} w_i \times K_i + w_m \times G$$

25 and sends it to the verifier as part of the protocol in [Bootle 2016].

5. The prover also sends the *key opening* for the vector commit:

$$ko_m = \rho \times F + \sum_{i=1}^{m-1} w_i \times K_i$$

6. The verifier calculates a “public key opening” of the key-statement wire (wire key), via elliptic curve arithmetic:

$$pk_m = Com(\mathbf{w}) - ko_m$$

5

Proof of equivalence of a hash pre-image and elliptic curve private key

This section described an example of a *key-statement zero-knowledge proof* that can be utilised in the context of hash attach bounties.

10

By way of example only, the statement $\mathcal{S}$ to be proved in order to claim the bounty is specifically formulated in this example as:

$\mathcal{S}$: “Given a SHA-256 hash function (H) with a public output h and a public point R on the secp256k1 elliptic curve, the secret pre-image of the hash, d (i.e. $h = H(d)$) is equal to the elliptic curve point multiplier (i.e. the corresponding ephemeral key, i.e. $R = d \times G$)”

In the approach described above, this statement consists of a single arithmetic circuit for the SHA-256 hash function $\mathcal{C}_{SHA256}$ (with m wires w_i ($i = 1, \dots, m$) and q gates) along with an assertion that the input wire (w_1) is the private key for public point R and that the output wire (w_m) is equal to h . i.e.:

20

$$\mathcal{S}: \mathcal{C}_{SHA256} \text{ is satisfied by the wires } \{w_i\}_{i=1}^m \text{ AND } w_1 \times G = R \text{ AND } w_m = h$$

Therefore, to fully verify this statement, the prover must demonstrate to the verifier that they know a satisfying assignment to the SHA256 circuit using the secp256k1 based commitment scheme, and then simply provide the *key-opening* for wire 1 (ko_1) and the *full opening* for wire m (w_m, ρ_m). The verifier does not learn the value of the input wire (w_1), or the values of any of the other wires except for the fully opened output wire w_n .

25

30

A proof verification algorithm $\mathcal{V}$ can therefore be constructed in an r-puzzle framework as follows.

In order to verify statement $\mathcal{S}$ above, the proof verification algorithm requires the following:

- 5 E1) z_i – one for each additive gate and three for each multiplicative gate,
- E2) B_i – one for each additive gate,
- E3) W_i – three for each additive gate and three for each multiplicative gate,
- E4) e_i – two for each multiplicative gate,
- E5) C_i – three for each multiplicative gate ,
- 10 E6) ko_1
- E7) w_m and ρ_m

Together, E1-E7 constitute the proof π of the proof transaction 1404, and may be referred to herein as the elements of the proof π .

- 15 Note that ξ is intentionally omitted here. ξ is defined to be the challenge set by the verifier. In a non-interactive ZKP, ξ may be defined to be hash value of W_L, W_R , and B , or C_i . Therefore, the verifier will be able to work out ξ from the information given above. IN other words, it is sufficient for the challenge transaction 1402 to specify steps for deriving ξ , without having to specify an actual value of ξ . The verification algorithm $\mathcal{V}$ can, in turn,
- 20 derive the value of ξ from the proof π itself.

The verification algorithm $\mathcal{V}$ can then be implemented for a given public ephemeral key R as follows:

- 1. verify each gate of the arithmetic circuit – both additive (as set out in the Σ_{zero} section above) and multiplicative (as set out in the Σ_{prod} section above) – based on E1-E5 above,
- 25 2. verify the additional preimage requirements as follows:
 - a) Extract $ko_1 = \rho_1 \cdot F$
 - b) Calculate $\tilde{R} = W_1 - ko_1$
 - 30 c) Checks $\tilde{R} = R$, i.e. that the result of 2b) matches the public ephemeral key of the applicable signature,

- d) Check the equality $W_m = w_m \cdot G + \rho_m \cdot F$
- e) Check that w_m is indeed h .

Note: step e) is applicable where h is pre-specified. However, it can be omitted in the case
 5 of hash collision bounties, as will now be explained.

For hash collision bounties, two such proofs π_1, π_2 are received

In order for the prover to provide the required two proofs π_1, π_2 , two instances of the
 10 arithmetic circuit are constructed by the prover (first and second instances), in respect of
 colliding pre-images discovered by the prover.

Notationally, the elements of the first proof π_1 are E1-E7 written exactly as above; the
 equivalent elements of the second proof π' are denoted using the same characters, but
 15 primed:

- E1') z'_i – one for each additive gate and three for each multiplicative gate,
- E2') B'_i – one for each additive gate,
- E3') W'_i – three for each additive gate and three for each multiplicative gate,
- E4') e'_i – two for each multiplicative gate,
- 20 E5') C'_i – three for each multiplicative gate ,
- E6') ko'_1
- E7') w'_m and ρ'_m

The verification algorithm $\mathcal{V}$ can then be implemented for a given pair of public ephemeral
 25 keys R_1, R_2 as follows:

- 1A. Verify the first instance of the arithmetic circuit, based on E1-E5,
- 1B. Verify the second instance of the arithmetic circuit, based on E1'-E5',
- 2A. Verify the preimage requirements for π_1 :
 - a) Extract $ko_1 = \rho_1 \cdot F$
 - 30 b) Calculate $\tilde{R} = W_1 - ko_1$
 - c) Checks $\tilde{R} = R_1$

d) Check the equality $W_m = w_m \cdot G + \rho_m \cdot F$

2B. Verify the preimage requirements for π_2 :

a) Extract $ko'_1 = \rho'_1 \cdot F$

b) Calculate $\tilde{R}' = W'_1 - ko'_1$

5 c) Check $\tilde{R}' = R_2$,

d) Check the equality $W'_m = w'_m \cdot G + \rho'_m \cdot F$

3. Verify equality of preimages:

a) $w_m = w'_m$

10 Note, as is apparent from 3a), although h is shown separately from π_1 and π_2 of the proof transaction 1404 in Figure 14, it is not in fact required as a separate element, and can be defined implicitly by $w_m = w'_m$.

The input wire w_1, w_1' are withheld, and those values are not derivable from the provided
15 key openings ko_1, ko'_1 on the input wires w_1, w_1' . In this manner, the key opening(s) is revealed without revealing the input wire(s).

ALTERNATIVE IMPLEMENTATION IN AN ACCOUNT-BASED MODEL

20 The above has largely been described in terms of an implementation in an output-based model (e.g. UTXO-based model). However it will be appreciated this is not limiting. Figure 11 illustrates a possible alternative implementation using an account-based model.

In short, in an account-based model, the r-puzzle functionality can be included in a smart
25 contract function which is called by the users. One party can set the r-puzzle value (or hashed r-puzzle value) in the smart contract and then the other party would then provide the signature to the smart contract afterwards.

In a UTXO blockchain architecture, the requirements embodied in the unlocking script of a
30 first transaction must be fulfilled by the locking script of a second transaction in order for the second transaction to be accepted as valid and recorded in the blockchain. In the

present context, this is beneficial as it leverages the work which has already been done by miners as part of the transaction validation process. As a concrete example in the present context, the fact that a transaction has been added to the blockchain implies that it has been verified by nodes throughout the blockchain network, which in turn implies that its locking script fulfils some specific useful requirements. An interested party does not need to check for themselves whether those requirements are satisfied – they can simply assume those requirements to be satisfied by virtue of the fact that the transaction has been recorded successfully in the blockchain. This stems from the fact that the script must return a result of “true” on completion in order for the transaction to be valid (there may be other requirements for the transaction to be valid), and if the script returns a result of “false” (which, according to the terminology used herein, includes the case where the script fails e.g. because an `OP_VERIFY` opcode terminates the script), the transaction is invalid.

However, in other blockchain models (e.g. certain account-based architectures), this interdependency between transaction validity and the result of running transaction code is not necessarily mirrored. For example, in certain smart contract blockchains, transactions may be valid and thus accepted for recordal on the blockchain provided they meet a set of “basic” validity requirements imposed by the blockchain protocol. Hence the second transaction may still be accepted as valid and recorded in the blockchain even if it does not satisfy some requirement embodied in the code of the first transaction. The code of the first transaction may for example be smart contract code.

Assuming the second transaction is addressed to a smart contract account created by the first transaction, it is then down to the smart contract code to determine how to respond to that transaction – it could for example ignore it (or otherwise return a result of false) if some requirement is not met, whereas if that requirement is correct, it can reward the prover with a quantity of digital asset deducted from the balance of the smart contract account and credited (or otherwise return a result of true). In a sense, this abstracts “agent-level” processing by smart contracts (agents), i.e. coded explicitly in smart contract code, from “protocol-level” processing performed “implicitly” by nodes i.e. processing that is performed on a transaction determine whether it meets the requirements of validity

imposed by a blockchain protocol on which the blockchain network operates. Hence, in such blockchain architectures, a “valid/invalid” decision by a node at the protocol level in respect of a transaction may be decoupled from a “true/false” result returned in respect of that transaction at the agent level by a smart contract, in that a transaction may be
5 determined to be valid at the protocol level but nonetheless return a result of false at the agent level.

This is in contrast to a UTXO architecture, in which the script returning a result of “true” is required for the transaction to be valid; the transaction is invalid if the script terminates or
10 completes leaving anything other than true on the stack (either of those outcomes constitutes a result of “false” as that term is used herein).

One of the basic requirements for transaction validity may be that the transaction includes a valid signature. Hence, whilst in the above UTXO examples, a signature is verified by the
15 code of a challenge transaction itself (e.g. using a `OP_CHECKSIG` opcode which verified the signature and returns true/false for the signature verification or an `OP_CHECKSIGVERIFY` opcode which checks the signature in the same way and additionally verifies the result is true, with the script terminating if it is not), in alternative blockchain architectures the signature may be verified by the processing node implicitly in the above sense, which may
20 avoid the need to code the signature check in the transaction code itself.

As a concrete example in the present context, a transaction may be deemed valid at the protocol level e.g. because it comprises a valid signature, but still return a result of false at the application-level e.g. because some other requirement is not satisfied.

25

Figure 11 shows an alternative of the node software 400 for processing transactions in accordance with an account-based model, the node software being labelled here 400acc. An instance of this node software 400acc may be implemented at each of the nodes 104 of an account-based version of the network 106. The account-based node software 400acc
30 comprises an account-based protocol engine 401acc, a contract engine 402acc (somewhat analogous to the script engine 402), an application-level decision engine 404, and a set of

one or more blockchain-related functional modules 405. At any given node 104, these may include any one, two or all three of: a mining module 405M, a forwarding module 405F and a storing module 405S (depending on the role or roles of the node). The protocol engine 401acc is configured to recognize the different fields of a transaction and process them in accordance with the node protocol. The node software 400acc also maintains an account state 406 of each of a plurality of accounts in memory of the respective node 104. These could for example include the accounts of Alice, the prover (e.g. Bob), and/or another party to be debited or credited depending on the contract to be enacted between Alice and the prover. The contract engine 402acc is arranged to modify the account state depending on the outcome of a smart contract received in a transaction. Smart contracts are also referred to as “agents”.

Figure 11 also shows a pair of transactions Tx_1^{acc} and Tx_2^{acc} which may implement the same or similar r-puzzle functionality as described above in relation to Figures 7 to 10. Each comprises a source account address 1102 (in a source address field) and a destination account address 1103 (in a destination address field). The first transaction Tx_1^{acc} comprises a source account address 1102a and a destination account address 1103a. The second transaction Tx_2^{acc} comprises a source account address 1102b and a destination account address 1103b. The first transaction Tx_1^{acc} also comprises a smart contract 1101. The smart contract 1101 may include a challenge (puzzle) by Alice. It may be created by Alice or by a third party on behalf of Alice using details provided by Alice. The second transaction Tx_2^{acc} may optionally comprises one or more free data fields 1104 for carrying user-specified payload data. This/these may comprise at least part of the solution to the puzzle provided by the prover, e.g. Bob. The transactions Tx_1^{acc} and Tx_2^{acc} are also signed by Alice and the prover respectively. Each transaction also comprises a signature 1105a, 1105b of the respective party.

The transactions are broadcast over the network 106. When the protocol engine 401acc receives each transaction it implicitly verifies whether or not the signature 1105 is valid. I.e. this is an inherent feature of the protocol engine 401acc and does not need to be specified in the smart contract 1101. The protocol engine 401acc thus validates each transaction for

forwarding and/or mining at least on condition that the respective signature is valid. It may also require one or more additional conditions for validity to be met. If valid, the application-level decision engine 404 can select whether to control the mining module 405M and/or forwarding module 405F to mine and/or forward the transaction respectively.

5

In such an account-based model, Alice, Bob and the smart contract itself are assigned separate accounts, with different account addresses. A transaction is said to be sent “to” the address in its destination address field, “from” the address in its source address field. In order to create an account for a smart contract, a transaction containing bytecode for the smart contract is uploaded to the blockchain in a transaction. For such an account-creating transaction, the destination address 1103 in the destination field should be an address which has never been used before in the blockchain, and once the transaction has been accepted, that address becomes the address of the newly-created smart contract account. Thereafter, a further transaction can be sent to that address in order to “call” the smart contract, i.e. cause the bytecode of the smart contract to be run in dependence on the further transaction. The “destination” address 1103 acts as an intermediary address for enacting the contract – Alice sends TX_1^{acc} to that address to create the smart contract which specifies one or more requirements; Bob sends TX_2^{acc} to that same address in order to call the smart contract, which in turn causes the smart contract to verify whether or not TX_2^{acc} satisfies those specified requirements. The “source” address 1102 specifies an account of a user that is party to the contract – in the case that the smart contract determines that TX_2^{acc} does meet the specified requirements, the smart contract can be configured to deduct an amount of digital asset from its own account balance, and cause the balance of the account having the source address 1102b in TX_2^{acc} (i.e. Bob’s account) to be credited by that amount (intuitively, by sending TX_2^{acc} , Bob effectively asks the smart contract (as identified in the destination address field) to credit his account (as identified in the source address field).

When the protocol engine 401acc receives X_2^{acc} , then on condition that it is valid, it will look for an account which matches the destination address 1103b in TX_2^{acc} . Assuming TX_1^{acc} has been processed and is valid, that account will exist by virtue of TX_1^{acc} and will be

30

associated with the smart contract code provided in TX_1 . In response, the protocol engine 401acc controls the contract engine 402acc to run the smart contract 1101 from Tx_1^{acc} , taking data from one or more fields of the smart contract as operand data, depending on what criteria are defined in the contract. The operand data may for example comprise the data from one or more of the free data fields 1104, and/or the signature from the signature field 1105b. On condition that the operand data from Tx_2^{acc} meets the one or more criteria defined in the smart contract 1101 of Tx_1^{acc} , then the contract engine 402acc modifies the account state 406 of one or more parties (Alice, the prover and/or one or more third parties) in accordance with a modification defined in the smart contract 1101. Otherwise this modification to the account state 406 is not made. Note however that in some account-based systems, the outcome of the smart contract is not a condition for validity of transactions. Hence if Tx_2^{acc} fails to meet the criteria set in the smart contract 1101 of Tx_1^{acc} , then Tx_2^{acc} will still be propagated and mined into a block as a record of the failed transaction. It may also still effect a mining fee (hence the protocol engine 401 may still modify the account state 406 of one of the parties and the winning miner).

To implement an r-puzzle, at least some of the r-puzzle functionality can be coded into the smart contract 1101 of Tx_1^{acc} , and the solution can be presented in one or more of the data fields 1104 of Tx_2^{acc} . For example this could be used to implement the variant of Figure 7. Optionally, some of the implicit signature verification functionality of the protocol engine 401acc could be exploited, e.g. to implement the one of the variants of Figures 8 to 10. In the case of Figures 8-10, steps II) and III) may be implicit functions of the protocol engine 401acc when it verifies the signature of Tx_2^{acc} (remember that signature verification per se is an inherent feature of the node protocol implemented by the protocol engine 401acc). Hence it is only required to layer step I) on top of this in the smart contract 1101 of Tx_1^{acc} . The smart contract checks whether the outcome of I) is true and whether the protocol engine 401ac indicates that Tx_2^{acc} is valid. If yes to both, then it declares an overall result of "true" for the verification, i.e. Bob has successfully met the challenge set by the r-puzzle. Note that of implementations of Figures 8-10, only the data value d in the cases of Figure 9 and 10 needs to be included in a free data field 1104. The signature information is included in the signature field 1105b.

A smart contract account also has indexed “data registers” (not shown) which are (logical) data storage elements associated with the account. In the above-outlined UTXO model, values are embedded in the locking script itself, and the same may be true of a particular
 5 piece of smart contract code 1101. However, smart contract bytecode of a smart contract may alternatively or additionally be run on data stored in one or more of its account registers. Moreover, it is generally possible to store values in the smart contract account registers after the smart contract account has been created. So, for example, a smart contract account may be created by a challenge transaction $Tx_{1,\alpha}^{acc}$ which contains the smart
 10 contract bytecode. A separate “intermediate” transaction $Tx_{1,\beta}^{acc}$ may then be sent to the (now existent) smart contract account, which has the effect of storing a particular value v in a register $\$R$ of the smart contract account. The smart contract may be configured to only accept such data from a specified source account address (for example), e.g. the same party who created the smart contract in the first place (Alice). When Tx_2^{acc} is received, the
 15 operations performed by the contract engine 402acc (e.g. “access register $\$R$ and compare the value to the value in data field $\$D$ of Tx_2^{acc} ”) are defined by the smart contract bytecode provided in the challenge transaction $Tx_{1,\alpha}^{acc}$; but the value stored in $\$R$ has been set by the intermediate transaction $Tx_{1,\beta}^{acc}$. According to the terminology used herein, $Tx_{1,\alpha}^{acc}$ is still said to be a challenge transaction which sets one or more requirements, only now
 20 those requirements may be defined with reference to data provided in one or more intermediate transactions (e.g. $Tx_{1,\beta}^{acc}$).

Thus, in some implementations, the challenge transaction $Tx_{1,\alpha}^{acc}$ may define the operations of the r-puzzle (e.g. compare the r-part of the signature of the proof transaction Tx_2^{acc} with
 25 the value in register $\$R$ to see if they match etc.) but the value in $\$R$ that is compared to the r-part of the proof transaction Tx_2^{acc} may have been set by the intermediate transaction $Tx_{1,\beta}^{acc}$.

Note also: some account-based models do not require the public key P to be included with
 30 the signature 1105. Instead they simply include a 1-bit flag, flg . As mentioned, it is possible to derive two possible keys P and $-P$ from (r, s) and the message. The flag flg is used to

signal which of these two possible solutions is in fact the public key corresponding to the private key V used by the prover to sign the message in Tx_2^{acc} . The protocol engine 401acc this uses (r, s) and flg to derive the prover's public key P instead of receiving it explicitly in Tx_2^{acc} . This technique would also be possible in output-based models and is not specific to
5 account-based models, but in the scripting languages used in many current output-based models there happens to be no dedicated opcode for deriving P from r and s , so it would be complex to code this functionality explicitly into the unlocking script using existing, general-purpose opcodes of stack-based language. It is further noted that, certain account-based models derive the source address of the transaction from the public key used to sign that
10 transaction. Hence, the source address is not necessarily encoded separately in the transaction, and in the case that the public key is derived from the signature, this means that the source address may also be indirectly derived from the signature.

It will be appreciated that the above embodiments have been described by way of example
15 only.

More generally, according to one aspect disclosed herein ("Example 1") there is provided a computer-implemented method of processing at least one proof transaction at a node of a blockchain network. The at least one proof transaction is received at the node and
20 comprises at least one Elliptic Curve Digital Signature Algorithm (ECDSA) signature and at least one zero-knowledge proof (ZKP) component. The node verifies the ECDSA signature of the proof transaction based on a public key associated with the ECDSA signature and a signed part of the transaction, and determined whether the ZKP component is correct for the ECDSA signature and a defined hash value and a defined hash function, in that it proves
25 an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of the defined hash value with respect to the defined hash function.

Example embodiments are set out below as further enumerated examples.

30 Example 2: An embodiment of the method of Example 1, wherein the at least one proof transaction comprises a second ZKP component and a second ECDSA signature, wherein

node verifies the second ECDSA signature, and determines whether the first and second ZKP components together prove knowledge of a successful collision attack on the defined hash function by:

- checking the first and second ECDSA signatures have different r-parts,
- 5 determining whether the first ZKP component is correct for the ECDSA signature and the defined hash value and the defined hash function, and
- determining whether the second ZKP component is correct for the second ECDSA signature and the same defined hash value and the same defined hash function as the first ZKP component, in that it proves an ephemeral key counterpart to an r-part of the second
- 10 ECDSA signature to correspond to a preimage of that defined hash value with respect to that defined hash function.

Example 3. An embodiment of the method of Example 2, wherein the second ECDSA signature is verified based on the same public key as ECDSA signature.

15

Example 4. An embodiment of the method of Example 1 or 2, wherein the at least one proof transaction comprises an anti-forgery signature, and the node verifies the anti-forgery signature based on the same public key as the ECDSA signature.

- 20 Example 5. An embodiment of the method of Example 1, 2, 3 or 4, wherein the at least one proof transaction identifies code of at least one challenge transaction, and thereby causes the node to run the code of the at least one challenge transaction in dependence on the at least one proof transaction, which in turn causes the node to determine whether the or each ZKP component of the at least one proof transaction is correct by determining
- 25 whether that ZKP component meets one or more proof requirements embodied in the code of the at least one challenge transaction.

- Example 6. An embodiment of the method of Example 5, wherein the public key used to verify the ECDSA signature is indicated in the at least one proof transaction but is not
- 30 specified by the at least one challenge transaction.

Example 7. An embodiment of the method of Example 6, wherein the public key is encoded as a string in the at least one proof transaction, and thereby indicated in the at least one proof transaction, or derived from the ECDSA signature of the at least one proof transaction, whereby the public key is indicated by the ECDSA signature itself.

5

Example 8. An embodiment of the method of any of Examples 5 to 7, wherein the hash value is defined in the at least one proof transaction and is not specified by the at least one challenge transaction.

10 Example 9. An embodiment of the method of any of Examples 5 to 7, wherein the hash value is defined in the at least one challenge transaction.

Example 10. An embodiment of the method of any of Examples 5 to 9 when dependent on Example 2, wherein the code of the at least one challenge transaction additionally embodies
15 a requirement for the first and second ECDSA signatures to have different r-parts.

Example 11. An embodiment of the method of any of Examples 5 to 9 when dependent on Example 3, wherein the code of the at least one challenge transaction additionally embodies a requirement for the first and second ECDSA signatures to have been generated using the
20 same private key.

Example 12. An embodiment of the method of any of Examples 5 to 11, wherein the code of the at least one challenge transaction causes the node to verify the ECDSA signature or each ECDSA signature in the at least one proof transaction.

25

Example 13. An embodiment of the method of any of Examples 5 to 11, wherein the code of the challenge transaction comprises locking script of one or more unspent transaction outputs of the challenge transaction, and the ECDSA signature and the ZKP component are embedded in unlocking script of the proof transaction.

30

Example 14. An embodiment of the method of Example 13, wherein the code of the challenge transaction is run in dependence on the proof transaction by concatenating the locking script of the challenge transaction with the unlocking script of the proof transaction, and running the concatenated script.

5

Example 15. An embodiment of the method of any of Examples 5 to 12, wherein the code of the at least one challenge transaction is virtual machine bytecode, and the or each ZKP component is provided in one or more data fields of the at least one proof transaction.

10 Example 16. An embodiment of the method of any of Examples 5 to 11, wherein the verification of the ECDSA signature is an implicit function performed by the node in accordance with a blockchain protocol of the blockchain network.

15 Example 17. An embodiment of the method of any of Examples 5 to 16, wherein the code of the at least one challenge transaction returns a result of false if the ZKP component is incorrect, and wherein the code requires the ZKP component to be correct in order to return a result of true.

20 Example 18. An embodiment of the method of Example 17, wherein the code also returns the result of false if the verification of the ECDSA signature fails, wherein the code requires both the ZKP component to be correct and the ECDSA signature to be successfully verified in order to return the result of true.

25 Example 19. An embodiment of the method of Example 17 or 18 when dependent on Example 2, wherein the result of false is returned if the verification of any of the ECDSA signatures fails and/or if any of the ZKP components is incorrect and/or if the first and second ECDSA signatures have the same r-part, and wherein, in order to return the result of true, each of the ECDSA signatures is required to be successfully verified, each of the ZKP components is required to be correct, and the first and second ECDSA signatures are
30 required to have different r-parts.

Example 20. An embodiment of the method of any of Examples 17 to 19 when dependent on Example 4, wherein the anti-forgery signature is required to be successfully verified in order to return the result of true, and the result of false is returned if the verification of the anti-forgery signature fails.

5

Example 21. An embodiment of the method of any of Examples 5 to 20, wherein the at least one proof transaction comprises a transaction identifier of the challenge transaction and thereby identifies the code of the challenge transaction.

10 Example 22. An embodiment of the method of any of Examples 5 to 20, wherein the at least one challenge transaction associates the code with an account address, and the at least one proof transaction comprises a matching account address and thereby identifies the code of the challenge transaction.

15 Example 23. An embodiment of the method of any of Examples 5 to 22, wherein the hash function is defined in the at least one challenge transaction.

Example 24. An embodiment of the method of any preceding Example, wherein the processing steps are performed in order to validate the proof transaction, and if the proof
20 transaction is determined to be valid, the node causes the proof transaction to be recorded in a blockchain maintained by the blockchain network.

Example 25. An embodiment of the method of Example 23 when dependent on any of Examples 17 to 19, wherein the node requires the result of true to be returned by the code
25 of the challenge transaction in order to determine the transaction to be valid.

Example 26. A computer program comprising code which is configured so as, when run on one or processors of a blockchain network node, to cause the one or more processors to process a proof transaction according to any of Examples 1 to 25.

30

Example 27. A node for a blockchain network, the node comprising:

memory configured to store code; and
one or more processors coupled to the memory for running the code, wherein the code is configured, when run on the one or more processors to cause the one or more processors to process a proof transaction received at the network interface in accordance
5 with any of Examples 1 to 25.

Example 28. An embodiments of the node of Example 27, wherein the stored code comprises the code of the challenge transaction of Example 5 or any Example dependent thereon.
10

Example 29. At least one challenge transaction comprising code which is configured, when run on a blockchain network node in dependence on at least one proof transaction, to cause the blockchain network node to determine whether a zero-knowledge proof (ZKP) component of the at least one proof transaction is correct for an Elliptic Curve Digital
15 Signature Algorithm (ECDSA) signature of the at least one proof transaction and a defined hash value and a defined hash function, in that it proves an ephemeral key counterpart to an r-part of the ECDSA signature to be a preimage of the defined hash value with respect to the defined hash function.

20 Example 30. An embodiment of the at least one challenge transaction of Example 29, wherein the code is configured to cause the blockchain network node to verify the ECDSA signature of the at least one proof transaction based on a public key associated with the ECDSA signature and a signed part of the at least one proof transaction.

25 Example 31. A method comprising the steps of creating at least one challenge transaction in accordance with Examples 29 or 30 and submitting the at least one challenge transaction to a node of a blockchain network for processing.

Example 32. A transaction generating computer system comprising:
30 a network interface;
memory configured to store code; and

one or more processors coupled to the memory for running the code, wherein the code is configured is configured, when run on the one or more processors, to cause the one or more processors to create at least one challenge transaction in accordance with Example 29 or 30, and submit the at least one challenge transaction via the network interface to a
5 node of a blockchain network for processing.

Example 33. At least one proof transaction for recoding in a blockchain, the at least one proof being embodied on a computer-readable data medium or media and comprising an Elliptic Curve Digital Signature Algorithm (ECDSA) signature, a signed part, and at least one
10 zero-knowledge proof (ZKP) component which proves an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of a defined hash value with respect to a defined hash function, without revealing the preimage in the at least one proof transaction.

15 Example 34. An embodiment of the proof transaction of Example 33, comprising at least two ZKP components and at least two ECDSA signatures having different r-parts, wherein the least two ZKP components prove the different r-parts to correspond to colliding preimages with respect to the defined hash function, without revealing the colliding preimages.

20

Example 35. An embodiment of the method, transaction, node or system of any preceding Example, wherein the or each ZKP component pertains to an arithmetic circuit, and proves the hash function to be satisfied by a set of circuit wires of the arithmetic circuit, and reveals (i) an output wire of the arithmetic circuit and (ii) a key opening on an input wire of the
25 arithmetic circuit without revealing the input wire.

Example 36. An embodiment of the method, transaction, node or system of Example 35, wherein the node verifies the arithmetic circuit based on the ZKP component, and uses the key opening to check that a wire key associated with the input wire matches the r-part of
30 the ECDSA signature.

Example 37. An embodiment of the method, transaction, node or system of any of Examples 35 to 36, wherein the hash value is defined by the output wire.

Example 38. An embodiment of the method of any of examples 1 to 27 or 35 to 37 26, the
5 method, comprising: using the at least one ZKP component to identify a vulnerability in the
blockchain network or another cryptographic system that is secured based on the hash
function, and taking one or more actions to mitigate the identified vulnerability.

Example 39. The method of Example 38, wherein the one or more actions comprise one or
10 more of the following actions:

revoking a digital certificate reliant on the hash function,

redeeming one or more unspent outputs of a blockchain transaction, the one or
more unspent outputs locked based on the hash function, to prevent a malicious party from
exploiting the attack to redeem the one or more unspent outputs.

15

For each of a plurality of transactions including the target transaction, at least some nodes
of the network may be configured to propagate each transaction on condition of the
transaction being valid and at least some nodes are configured to record each transaction in
the copy of the blockchain at that node on condition of the transaction being valid. The
20 validity of the target transaction is conditional on at least the ECDSA signature being valid
and the ZKP component being correct.

In embodiments, any of the above transaction(s) may be processed by a node in order to
validate the proof transaction, and if the proof transaction is determined to be valid, the
25 node causes the proof transaction to be recorded in a blockchain maintained by the
blockchain network.

For example, such validation may be applied in a UTXO model.

Alternatively or additionally, any of the above transaction(s) may be processed by a node which returns one of a result of true and a result of false (and a result of true may or may not be required in that event for the transaction(s) to be valid).

- 5 For example, in an account-based model, a valid transaction may nonetheless return the result of false.

- In any r-puzzle context, it may be the case that a public key used to verify an ECDSA signature of at least one proof transaction is indicated in the at least one proof transaction
- 10 but is not specified by the at least one challenge transaction (or elsewhere on the blockchain or otherwise). Accordingly any private key may be used to generate the ECDSA signature (hence the signature may be valid irrespective of whose private key is used to generate it).
- 15 The public key may be encoded as a string in the at least one proof transaction, and thereby indicated in the at least one proof transaction, or derived from the ECDSA signature of the at least one proof transaction, whereby the public key is indicated by the ECDSA signature itself.
- 20 The at least one proof transaction may comprise a transaction identifier of the challenge transaction and thereby indicate the challenge transaction (or an applicable component thereof, such as an r-puzzle, code etc.).

- Alternatively, the at least one challenge transaction may associate an r-puzzle, code or other
- 25 component with an account address, and the at least one proof transaction may comprise a matching account address and thereby indicate that component of the challenge transaction.

- According to another aspect disclosed herein, there may be provided a method comprising
- 30 the actions of the first party (challenger), second party (prover), any third party that may be involved, and the network of nodes (blockchain network).

According to another aspect disclosed herein, there may be provided a system comprising the computer equipment of the first party, the computer equipment of the second party, the computer equipment of any third party, and the network of nodes.

5

Other variants or use cases of the disclosed techniques may become apparent to the person skilled in the art once given the disclosure herein. The scope of the disclosure is not limited by the described embodiments but only by the accompanying claims.

Claims

1. A computer-implemented method of processing at least one proof transaction at a node of a blockchain network, the method comprising:
 - receiving the at least one proof transaction, wherein the at least one proof transaction comprises at least one Elliptic Curve Digital Signature Algorithm (ECDSA) signature and at least one zero-knowledge proof (ZKP) component;
 - verifying the ECDSA signature of the at least one proof transaction based on a public key associated with the ECDSA signature and a signed part of the at least one proof transaction; and
 - determining whether at least one ZKP component proves knowledge of a successful attack on the hash function by determining whether the ZKP component is correct for the ECDSA signature and a defined hash value and a defined hash function, in that it proves an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of the defined hash value with respect to the defined hash function.
2. The method of claim 1, wherein the at least one proof transaction comprises a second ZKP component and a second ECDSA signature, wherein node verifies the second ECDSA signature, and determines whether the first and second ZKP components together prove knowledge of a successful collision attack on the defined hash function by:
 - checking the first and second ECDSA signatures have different r-parts,
 - determining whether the first ZKP component is correct for the ECDSA signature and the defined hash value and the defined hash function, and
 - determining whether the second ZKP component is correct for the second ECDSA signature and the same defined hash value and the same defined hash function as the first ZKP component, in that it proves an ephemeral key counterpart to an r-part of the second ECDSA signature to correspond to a preimage of that defined hash value with respect to that defined hash function.
3. The method of claim 2, wherein the second ECDSA signature is verified based on the same public key as ECDSA signature.

4. The method of claim 1 or 2, wherein the at least one proof transaction comprises an anti-forgery signature, and the node verifies the anti-forgery signature based on the same public key as the ECDSA signature.
5. The method of claim 1, 2, 3 or 4, wherein the at least one proof transaction identifies code of at least one challenge transaction, and thereby causes the node to run the code of the at least one challenge transaction in dependence on the at least one proof transaction, which in turn causes the node to determine whether the or each ZKP component of the at least one proof transaction is correct by determining whether that ZKP component meets one or more proof requirements embodied in the code of the at least one challenge transaction.
6. The method of claim 5, wherein the public key used to verify the ECDSA signature is indicated in the at least one proof transaction but is not specified by the at least one challenge transaction.
7. The method of claim 6, wherein the public key is encoded as a string in the at least one proof transaction, and thereby indicated in the at least one proof transaction, or derived from the ECDSA signature of the at least one proof transaction, whereby the public key is indicated by the ECDSA signature itself.
8. The method of any of claims 5 to 7, wherein the hash value is defined in the at least one proof transaction and is not specified by the at least one challenge transaction.
9. The method of any of claims 5 to 7, wherein the hash value is defined in the at least one challenge transaction.
10. The method of any of claims 5 to 9 when dependent on claim 2, wherein the code of the at least one challenge transaction additionally embodies a requirement for the first and second ECDSA signatures to have different r-parts.

11. The method of any of claims 5 to 9 when dependent on claim 3, wherein the code of the at least one challenge transaction additionally embodies a requirement for the first and second ECDSA signatures to have been generated using the same private key.
12. The method of any of claims 5 to 11, wherein the code of the at least one challenge transaction causes the node to verify the ECDSA signature or each ECDSA signature in the at least one proof transaction.
13. The method of any of claims 5 to 11, wherein the code of the challenge transaction comprises locking script of one or more unspent transaction outputs of the challenge transaction, and the ECDSA signature and the ZKP component are embedded in unlocking script of the proof transaction.
14. The method of claim 13, wherein the code of the challenge transaction is run in dependence on the proof transaction by concatenating the locking script of the challenge transaction with the unlocking script of the proof transaction, and running the concatenated script.
15. The method of any of claims 5 to 12, wherein the code of the at least one challenge transaction is virtual machine bytecode, and the or each ZKP component is provided in one or more data fields of the at least one proof transaction.
16. The method of any of claims 5 to 11, wherein the verification of the ECDSA signature is an implicit function performed by the node in accordance with a blockchain protocol of the blockchain network.
17. The method of any of claims 5 to 16, wherein the code of the at least one challenge transaction returns a result of false if the ZKP component is incorrect, and wherein the code requires the ZKP component to be correct in order to return a result of true.

18. The method of claim 17, wherein the code also returns the result of false if the verification of the ECDSA signature fails, wherein the code requires both the ZKP component to be correct and the ECDSA signature to be successfully verified in order to return the result of true.
19. The method of claim 17 or 18 when dependent on claim 2, wherein the result of false is returned if the verification of any of the ECDSA signatures fails and/or if any of the ZKP components is incorrect and/or if the first and second ECDSA signatures have the same r-part, and wherein, in order to return the result of true, each of the ECDSA signatures is required to be successfully verified, each of the ZKP components is required to be correct, and the first and second ECDSA signatures are required to have different r-parts.
20. The method of any of claims 17 to 19 when dependent on claim 4, wherein the anti-forgery signature is required to be successfully verified in order to return the result of true, and the result of false is returned if the verification of the anti-forgery signature fails.
21. The method of any of claims 5 to 20, wherein the at least one proof transaction comprises a transaction identifier of the challenge transaction and thereby identifies the code of the challenge transaction.
22. The method of any of claims 5 to 20, wherein the at least one challenge transaction associates the code with an account address, and the at least one proof transaction comprises a matching account address and thereby identifies the code of the challenge transaction.
23. The method of any of claims 5 to 22, wherein the hash function is defined in the at least one challenge transaction.
24. The method of any preceding claim, wherein the processing steps are performed in order to validate the proof transaction, and if the proof transaction is determined to be

valid, the node causes the proof transaction to be recorded in a blockchain maintained by the blockchain network.

25. The method of claim 23 when dependent on any of claims 17 to 19, wherein the node requires the result of true to be returned by the code of the challenge transaction in order to determine the transaction to be valid.

26. The method of any preceding claim, comprising:
using the at least one ZKP component to identify a vulnerability in the blockchain network or another cryptographic system that is secured based on the hash function, and
taking one or more actions to mitigate the identified vulnerability.

27. The method of claim 26, wherein the one or more actions comprise one or more of the following actions:
revoking a digital certificate reliant on the hash function,
redeeming one or more unspent outputs of a blockchain transaction, the one or more unspent outputs locked based on the hash function, to prevent a malicious party from exploiting the attack to redeem the one or more unspent outputs.

28. A computer program comprising code which is configured so as, when run on one or processors of a blockchain network node, to cause the one or more processors to process a proof transaction according to any of claims 1 to 27.

29. A node for a blockchain network, the node comprising:
memory configured to store code; and
one or more processors coupled to the memory for running the code, wherein the code is configured, when run on the one or more processors to cause the one or more processors to process a proof transaction received at the network interface in accordance with any of claims 1 to 27.

30. The node of claim 29, wherein the stored code comprises the code of the challenge transaction of claim 5 or any claim dependent thereon.

31. At least one challenge transaction comprising code which is configured, when run on a blockchain network node in dependence on at least one proof transaction, to cause the blockchain network node to determine whether a zero-knowledge proof (ZKP) component of the at least one proof transaction is correct for an Elliptic Curve Digital Signature Algorithm (ECDSA) signature of the at least one proof transaction and a defined hash value and a defined hash function, in that it proves an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of the defined hash value with respect to the defined hash function.

32. The at least one challenge transaction of claim 31, wherein the code is configured to cause the blockchain network node to verify the ECDSA signature of the at least one proof transaction based on a public key associated with the ECDSA signature and a signed part of the at least one proof transaction.

33. A method comprising the steps of creating at least one challenge transaction in accordance with claims 31 or 32 and submitting the at least one challenge transaction to a node of a blockchain network for processing.

34. A transaction generating computer system comprising:
a network interface;
memory configured to store code; and
one or more processors coupled to the memory for running the code, wherein the code is configured is configured, when run on the one or more processors, to cause the one or more processors to create at least one challenge transaction in accordance with claim 29 or 30, and submit the at least one challenge transaction via the network interface to a node of a blockchain network for processing.

35. At least one proof transaction for recoding in a blockchain, the at least one proof being embodied on a computer-readable data medium or media and comprising an Elliptic Curve Digital Signature Algorithm (ECDSA) signature, a signed part, and at least one zero-knowledge proof (ZKP) component which proves an ephemeral key counterpart to an r-part of the ECDSA signature to correspond to a preimage of a defined hash value with respect to a defined hash function, without revealing the preimage in the at least one proof transaction.

36. The proof transaction of claim 35, comprising at least two ZKP components and at least two ECDSA signatures having different r-parts, wherein the least two ZKP components prove the different r-parts to correspond to colliding preimages with respect to the defined hash function, without revealing the colliding preimages.

37. The method, transaction node or system of any preceding claim, wherein the or each ZKP component pertains to an arithmetic circuit, and proves the hash function to be satisfied by a set of circuit wires of the arithmetic circuit, and reveals (i) an output wire of set of circuit wires and (ii) a key opening on an input wire of the set of circuit wires without revealing the input wire.

38. The method, transaction, node or system of claim 37, wherein the node verifies the arithmetic circuit based on the ZKP component, and uses the key opening to check that a wire key associated with the input wire matches the r-part of the ECDSA signature.

39. The method, transaction, node or system of any of claims 37 to 38, wherein the hash value is defined by the output wire.