# Analysis of CryptoNote Transaction Graphs using the Dulmage-Mendelsohn Decomposition 

Saravanan Vijayakumaran<br>Department of Electrical Engineering Indian Institute of Technology Bombay<br>Mumbai, India<br>sarva@ee.iitb.ac.in


#### Abstract

Transactions in CryptoNote blockchains induce a bipartite graph, with the set of transaction outputs forming one vertex class and the set of key images forming the other vertex class. In this graph, an edge exists between an output and a key image if the output appeared in the ring of the linkable ring signature which created the key image. Any maximum matching on this graph is a plausible candidate for the ground truth, i.e. the association of each key image with the actual output being spent in the transaction. The Dulmage-Mendelsohn (DM) decomposition of a bipartite graph reveals constraints which are satisfied by every maximum matching on the graph. It identifies vertices which are matched in every maximum matching. It classifies edges as admissible or inadmissible. An edge is called admissible if it appears in at least one maximum matching and is called inadmissible if it does not appear in any maximum matching. The DM decomposition of a CryptoNote transaction graph reveals a set of outputs which can be marked as spent (precisely those outputs which are matched by every maximum matching). In some transaction rings, the decomposition identifies the true output being spent (making the ring traceable) by classifying the edges from all the other outputs to the key image as inadmissible. For pre-RingCT outputs in Monero, the DM decomposition performs better than existing techniques for Monero traceability, but the improvement is marginal. For RingCT outputs in Monero up to April 1, 2021, the DM decomposition is only able to identify the same five outputs that were identified as spent by existing techniques (which do not use information from hard forks).


Keywords: Cryptocurrency • CryptoNote • Monero • Traceability.

## 1 Introduction

Coins in CryptoNote blockchains are associated with stealth addresses, which are also called one-time addresses or transaction outputs 13. We will use the term output for brevity. Each output is uniquely identified by a public key, which is a point on an elliptic curve. To spend from an output, the spender needs to know the corresponding secret key. In a transaction, the spender creates a ring of outputs which is a set containing the output being spent and some other outputs sampled from the CryptoNote blockchain (these are called decoy outputs or mixins). The spender generates a linkable ring signature over the ring of outputs using the secret key of the output being spent. This signature only reveals that the signer knows the secret key corresponding to one of the ring outputs, without revealing the identity of the actual output being spent. To prevent double spending from an output, the linkable ring signature reveals the key image of the output being spent. The key image of an output is a unique deterministic function of the secret key. For example, in Monero the public key associated with an output is given by $P=x G$ where $G$ is the base point of the elliptic curve group used by Monero and $x$ is the secret key. Let $H_{\mathrm{p}}(\cdot)$ denote the Keccak hash function, whose outputs can be interpreted as points on the elliptic curve. The key image $I$ of the output associated with $P$ is given by $x H_{\mathrm{p}}(P)$. If the owner of the output corresponding to $P$ tries to spend the coins associated with it more than once, then the key image $I$ would appear again in the second transaction, identifying it as a double spending transaction. Such transactions are not included in blocks by miners as the resulting block would be considered invalid by the network.

Consider a CryptoNote transaction which spends from two existing outputs and creates three new outputs as illustrated in Fig. 1. The new outputs are denoted by $R_{1}, R_{2}, R_{3}$. The transaction has two rings of outputs of size five each, $\left(P_{1}, P_{2}, \ldots, P_{5}\right)$ and $\left(Q_{1}, Q_{2}, \ldots, Q_{5}\right)$. Exactly one output from each ring is being spent in the transaction. The key images $I_{1}$ and $I_{2}$ of the outputs being spent are revealed in the transaction. Note that the two rings can have common outputs.

For the purpose of illustration, suppose that the two rings have two outputs in common. Let $Q_{1}=P_{4}$ and $Q_{2}=P_{5}$. The relationship between the ring outputs and the key images in this transaction can be represented by the bipartite graph in Fig. 2. The union of the two ring output sets forms one vertex class and the two key images form the other vertex class. An edge between an output and a key image indicates that the latter could be the true key image of that output. Note that the new outputs $R_{1}, R_{2}, R_{3}$ play no role in the construction of the bipartite graph. In this document, we will refer to such output/key image bipartite graphs as transaction graphs.

As each key image must have been generated from a unique output, any pair of edges $\left(P_{i}, I_{1}\right)$ and $\left(Q_{j}, I_{2}\right)$ such that $P_{i} \neq Q_{j}$ is a plausible candidate for the true relationship between the outputs and key images. Recall that a matching on a graph is a subset of the edges such that no two edges in the subset share a vertex. The pair of edges $\left(P_{i}, I_{1}\right)$ and $\left(Q_{j}, I_{2}\right)$ with $P_{i} \neq Q_{j}$ is a matching on


Fig. 1. A CryptoNote transaction with two inputs and three outputs
the graph in Fig. 2. In fact, it is a matching of maximum size as any three edges in this graph would have two which meet in either $I_{1}$ or $I_{2}$.

Let us now consider a similar bipartite graph induced by the set of all transactions which have appeared up to the block having height $h$. The key image vertex class $\mathcal{K}_{h}$ in this graph is the set of all key images which have appeared on the blockchain up to block height $h$. The output vertex class $\mathcal{O}_{h}$ is the set of all outputs which have appeared in at least one transaction ring in the blocks up to height $h$. Note that $\mathcal{O}_{h}$ is not the set of all outputs which have appeared on the blockchain in blocks up to height $h$. For example, suppose that the transaction illustrated in Fig. 11 appeared on the blockchain in a block with height


Fig. 2. Transaction graph corresponding to the transaction in Fig. 1
$h^{\prime}<h$. Further, suppose that by the time the block with height $h$ appeared on the blockchain, the outputs $R_{1}$ and $R_{2}$ have appeared in transaction rings (as decoy or spending outputs) but $R_{3}$ has never appeared in a ring. Then the set $\mathcal{O}_{h}$ contains $R_{1}, R_{2}$ but not $R_{3}$. Thus $\mathcal{O}_{h}$ is the set of "ringed" outputs at block height $h$. It increases monotonically with $h$ as previously "unringed" outputs appear for the first time in transaction rings.

At block height $h$, we represent the edge set of the transaction graph induced by the CryptoNote transaction rings as a subset $E$ of $\mathcal{O}_{h} \times \mathcal{K}_{h}$. For $P \in \mathcal{O}_{h}$ and $I \in \mathcal{K}_{h}$, the edge $(P, I)$ belongs to $E$ if the output $P$ appeared in the transaction ring used to create $I$ (via the linkable ring signature).

Since each key image $I \in \mathcal{K}_{h}$ is generated from a unique output $P \in \mathcal{O}_{h}$, we have $\left|\mathcal{K}_{h}\right| \leq\left|\mathcal{O}_{h}\right|$. In a bipartite graph with vertex classes of cardinality $m$ and $n$, the size of a maximum matching can be at most $\min (m, n)$. Since the edges corresponding to the true association between outputs and key images form a matching of size $\left|\mathcal{K}_{h}\right|$, the induced bipartite graph always has a maximum matching. In fact, we have the following principle which has been discussed by Monero Research Lab researchers [6] and others [15].

Any maximum matching on the induced bipartite graph is a plausible candidate for the ground truth, i.e. the true association between outputs and key images.

We now describe the Dulmage-Mendelsohn decomposition of a bipartite graph and its relation to maximum matchings.

## 2 The Dulmage-Mendelsohn Decomposition

Consistent with notation used by Dulmage and Mendelsohn [5], we define an undirected bipartite graph $K$ as a triple $(S, T, E)$ where $S$ and $T$ are non-empty sets representing vertex classes and $E \subseteq S \times T$ represents the edge set. So an edge in $K$ is given by an ordered pair $(s, t)$ where $s \in S$ and $t \in T$. The ordering of the vertices in the edge $(s, t)$ is simply a consequence of putting $S$ before $T$ in the triple $(S, T, E)$, and does not imply directivity. We say that an edge $(s, t)$ belongs to the graph $K$, written as $(s, t) \in K$, to mean that $(s, t) \in E$. We will only consider bipartite graphs $K$ where both $S$ and $T$ are finite sets.

Definition 1. Let $K=(S, T, E)$ be a bipartite graph. Let $A$ and $B$ be subsets of $S$ and $T$ respectively. A pair of such sets $(A, B)$ is called a vertex cover for a bipartite graph $K$ if for each edge $(s, t) \in K$, either $s \in A$ or $t \in B$ (both conditions can also hold).

We state and prove a simple lemma for later reference.
Lemma 1. Suppose $(A, B)$ is a vertex cover of a bipartite graph $K=(S, T, E)$. Then $E \cap\left(A^{c} \times B^{c}\right)=\emptyset$.

Proof. We want to argue that the graph cannot have edges in the set $A^{c} \times B^{c}$. Suppose that $(s, t) \in E \cap\left(A^{c} \times B^{c}\right)$. Then $s \in A^{c}$ and $t \in B^{c}$. This contradicts the assumption that $(A, B)$ is a vertex cover of $K$.

Definition 2. The size of a vertex cover $(A, B)$ is defined as $|A|+|B|$ where $|X|$ denotes the cardinality of a set $X$.

Since $S$ and $T$ are assumed to be finite sets, every vertex cover of $K$ will have a finite size.

Definition 3. The cover number of a bipartite graph $K$ is the minimum of $|A|+|B|$ over all vertex covers $(A, B)$ of $K$.

Definition 4. A vertex cover $(A, B)$ of a bipartite graph $K$ whose size equals the cover number of $K$ is called a minimum cover.

We now define matchings on bipartite graphs and relate them to vertex covers. We say that edges $(s, t)$ and $\left(s^{\prime}, t^{\prime}\right)$ share a vertex if either $s=s^{\prime}$ or $t=t^{\prime}$.

Definition 5. A matching on a bipartite graph $K=(S, T, E)$ is a subset $M$ of the edge set $E$ such that no two edges in $M$ share a vertex. The set cardinality $|M|$ is called the order of the matching $M$.

Definition 6. A maximum matching on a bipartite graph $K$ is a matching on $K$ of maximum order.

The following theorem by König relates cover numbers to orders of maximum matchings.

Theorem 1. The cover number of a finite bipartite graph equals the order of maximum matchings on the graph.

The following definition classifies edges according to their membership in maximum matchings on $K$.

Definition 7. An edge $(s, t)$ of a bipartite graph $K$ is said to be admissible if there exists a maximum matching $M$ on $K$ such that $(s, t) \in M$. An edge which is not admissible is said to be inadmissible.

In the transaction graph induced by some CryptoNote transactions, if we can show that an edge $(P, I)$ is inadmissible, then $P$ cannot be the true output corresponding to the key image $I$. This fact reduces the effective ring size of the transaction which created $I$. If we can classify all the edges incident on $I$ except one as inadmissible, then the true output corresponding to $I$ is identified. We now state a central theorem (proved in [5]) which characterizes inadmissible edges in terms of minimum covers.

Theorem 2. An edge ( $s, t$ ) of a bipartite graph $K$ is inadmissible if and only if there exists a minimum cover $(A, B)$ of $K$ such that $(s, t)$ belongs to $A \times B$.

Example 1. To illustrate the above the theorem, consider three CryptoNote transaction rings having ring members $\left\{P_{1}\right\},\left\{P_{1}, P_{2}\right\}$, and $\left\{P_{1}, P_{2}, P_{3}, P_{4}\right\}$ respectively. Let $I_{1}, I_{2}, I_{3}$ be the key images created from these three transaction rings. The first transaction ring has only one member and therefore corresponds


Fig. 3. Transaction graph corresponding to Example 1
to a zero-mixin transaction. The induced bipartite graph is shown in Fig. 3. Let $S=\left\{P_{1}, P_{2}, P_{3}, P_{4}\right\}$ and $T=\left\{I_{1}, I_{2}, I_{3}\right\}$. To apply Theorem 2 , we need to find a minimum cover $(A, B)$ of the graph where $A \subseteq S$ and $B \subseteq T$. As any maximum matching on the graph has size 3 , minimum covers will also have size 3 .

- The set pair $(\emptyset, T)$ is a minimum cover of the graph. But it does not give us any inadmissible edges as $\emptyset \times T$ is empty.
- The set pair $\left(\left\{P_{1}\right\},\left\{I_{2}, I_{3}\right\}\right)$ is a minimum cover of the graph. Theorem 2 now tells us that the edges $\left(P_{1}, I_{2}\right)$ and $\left(P_{1}, I_{3}\right)$ are inadmissible. This is the same conclusion one draws from the fact that the ring which created $I_{1}$ is zero-mixin. Since $P_{1}$ is the true output corresponding to $I_{1}$, it cannot be the true output corresponding to $I_{2}$ or $I_{3}$.
- The set pair $\left(\left\{P_{1}, P_{2}\right\},\left\{I_{3}\right\}\right)$ is a minimum cover of the graph. Theorem 2 tells us that the edges $\left(P_{1}, I_{3}\right)$ and $\left(P_{2}, I_{3}\right)$ are inadmissible. This is the same conclusion one draws from the fact that the first two transaction rings $\left\{P_{1}\right\}$ and $\left\{P_{1}, P_{2}\right\}$ form a closed set (as defined in (15). Essentially, the outputs $P_{1}, P_{2}$ must be the true outputs corresponding to key images $I_{1}, I_{2}$. So neither of them can be the true output corresponding to $I_{3}$.

We will need the following corollary of Theorem 2 in a later argument.
Corollary 1. Let $K$ be a bipartite graph with finite cover number and let $(A, B)$ be a minimum cover of $K$. Every maximum matching $M$ on $K$ has $|A|$ edges in $A \times B^{c}$ and $|B|$ edges in $A^{c} \times B$.

Proof. By König's theorem (Theorem 11), the maximum matching $M$ has $|A|+|B|$ edges. Lemma 1 tells us that $M$ cannot have any edges in $A^{c} \times B^{c}$ and Theorem 2 tells us that $M$ has no edges in $A \times B$. Thus all the edges of $M$ must lie in either $A \times B^{c}$ or $A^{c} \times B$.

As distinct edges in the matching $M$ cannot share a vertex, for any two distinct edges $\left(s_{1}, t_{1}\right)$ and $\left(s_{2}, t_{2}\right)$ of $M$ in $A \times B^{c}$ we must have $s_{1} \neq s_{2}$. Thus the number of edges of $M$ in $A \times B^{c}$ is at most $|A|$. Similarly, the number of edges of $M$ in $A^{c} \times B$ is at most $|B|$. Since $M$ has exactly $|A|+|B|$ edges, the sets $A \times B^{c}$ and $A^{c} \times B$ must have exactly $|A|$ and $|B|$ edges of $M$, respectively.

Theorem 22 is also related to the definition of sets of spent outputs given by Monero Research Lab [11. We recall the definition below for convenience. Readers not interested in this relationship can skip ahead to Theorem 3 without loss of continuity.

Definition 8. Let $\mathcal{O}$ be the set of outputs on a CryptoNote-style blockchain. Let $R_{i} \subset \mathcal{O}$ be a transaction ring of outputs for $i=1,2, \ldots, n$. One output in each transaction ring is spent resulting in a unique key image. We say that each $R_{i}$ is spent if

$$
\left|\bigcup_{i=1}^{n} R_{i}\right|=n
$$

An output is spent if it is an element of a spent ring.
The reasoning behind this definition is as follows. Each ring $R_{i}$ has a unique key image $I_{i}$ associated with it. Since $\cup_{i=1}^{n} R_{i}$ has only $n$ outputs, all of them must have been spent to create the $n$ key images $I_{1}, I_{2}, \ldots, I_{n}$.

Let $\cup_{i=1}^{n} R_{i}=\left\{P_{1}, P_{2}, \ldots, P_{n}\right\}$. Suppose we draw the bipartite graph induced by the entire blockchain history while listing $P_{1}, \ldots, P_{n}$ and $I_{1}, \ldots, I_{n}$ before the other vertices on each side. Let $P_{n+1}, \ldots, P_{M}$ be the other outputs on the blockchain. Let $I_{n+1}, \ldots, I_{N}$ be the other key images on the blockchain where $N \leq M$. Fig. 4 illustrates the bipartite graph. Since each key image in $I_{1}, I_{2}, \ldots, \bar{I}_{N}$ corresponds to a unique true output on the left hand side, there exists a maximum matching of order $N$ on this graph. Then $\left(\emptyset,\left\{I_{1}, \ldots, I_{N}\right\}\right)$ is a minimum cover of the graph.

Note that there cannot be any edges from the key images $I_{1}, \ldots, I_{n}$ to the outputs $P_{n+1}, P_{n+2}, \ldots, P_{M}$. To see this, suppose there is an edge from $I_{j}$ to $P_{k}$ for some $j \in\{1,2, \ldots, n\}$ and $k \in\{n+1, \ldots, M\}$. Then $P_{k}$ must belong to the ring $R_{j}$ as it is the only ring which contributes edges incident on $I_{j}$. This would mean $P_{k}$ belongs to $\cup_{i=1} R_{i}=\left\{P_{1}, P_{2}, \ldots, P_{n}\right\}$, which is a contradiction as $k \geq n+1$. So all the edges incident on $I_{1}, \ldots, I_{n}$ must have an output from $P_{1}, \ldots, P_{n}$ on the other end.

The above argument shows that $\left(\left\{P_{1}, \ldots, P_{n}\right\},\left\{I_{n+1}, \ldots, I_{N}\right\}\right)$ is a minimum cover of the graph. Thus the union $\cup_{i=1}^{n} R_{i}$ of the spent rings as defined in Definition 8 is the first member of a minimum cover of the transaction graph. One can also prove the other direction. We claim that if $(A, B)$ is a minimum cover of the transaction graph where $A \neq \emptyset$, then there exist transaction rings $R_{i_{1}}, R_{i_{2}}, \ldots, R_{i_{n}}$ such that

$$
\begin{equation*}
A=\bigcup_{j=1}^{n} R_{i_{j}} \quad \text { and } \quad\left|\bigcup_{j=1}^{n} R_{i_{j}}\right|=n \tag{1}
\end{equation*}
$$

Let $\mathcal{O}=\left\{P_{1}, \ldots, P_{M}\right\}$ be the set of all outputs and $\mathcal{K}=\left\{I_{1}, \ldots, I_{N}\right\}$ be the set of all key images, which have appeared on the blockchain at some block height. For a minimum cover $(A, B)$, let $B^{c}=\mathcal{K} \backslash B$ be the set of key images not


Fig. 4. Transaction graph to illustrate the connection of Definition 8 to Theorem 2
in $B$. Suppose $B^{c}=\left\{I_{i_{1}}, I_{i_{2}}, \ldots, I_{i_{n}}\right\}$. Each key image $I_{i_{j}}$ in $B^{c}$ is associated with a unique transaction ring $R_{i_{j}}$ which contains the true output corresponding to it.

Since $(\emptyset, \mathcal{K})$ is a minimum cover of the graph, every minimum cover must have size $N$. This implies that $|A|+|B|=N$. As $n=\left|B^{c}\right|=N-|B|$, the set $A$ must have $n$ outputs.

Since $(A, B)$ is a cover of the bipartite graph, every edge incident on key images in $B^{c}$ must be covered by an output in $A$ (as $B$ can only cover edges incident on the key images in it). The ring $R_{i_{j}}$ associated with a key image $I_{i_{j}} \in B^{c}$ is the set of outputs adjacent to $I_{i_{j}}$ in the graph. So the other endpoints of edges incident on $I_{i_{j}}$ are in $R_{i_{j}}$. This implies that the transaction ring $R_{i_{j}}$ is a subset of $A$ for every $I_{i_{j}} \in B^{c}$. Thus $\cup_{j=1}^{n} R_{i_{j}} \subseteq A$.

Furthermore, $\left|\cup_{j=1}^{n} R_{i_{j}}\right| \geq n$ because each of the $n$ key images $I_{i_{1}}, I_{i_{2}}, \ldots, I_{i_{n}}$ has a unique true output in $\cup_{j=1}^{n} R_{i_{j}}$. Putting all this together, we have

$$
\begin{equation*}
n \leq\left|\bigcup_{j=1}^{n} R_{i_{j}}\right| \leq|A|=n \tag{2}
\end{equation*}
$$

Thus, we conclude that $A=\cup_{j=1}^{n} R_{i_{j}}$ and that $\left|\cup_{j=1}^{n} R_{i_{j}}\right|=n$. This completes our digression discussing the relationship between Definition 8 and Theorem 2 . We now return to our discussion of minimum covers.

The following two theorems were proved by Dulmage and Mendelsohn [5].

Theorem 3. If $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ are minimum covers of a bipartite graph $K$ having finite cover number, then $\left(A_{1} \cap A_{2}, B_{1} \cup B_{2}\right)$ and $\left(A_{1} \cup A_{2}, B_{1} \cap B_{2}\right)$ are both minimum covers of $K$.

Theorem 4. Let $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ be minimum covers of a bipartite graph $K$ having finite cover number. If $A_{1} \subseteq A_{2}$, then $B_{1} \supseteq B_{2}$.

Setting $A_{1}=A_{2}$ in the above theorem gives us the following corollary.
Corollary 2. If $\left(A, B_{1}\right)$ and $\left(A, B_{2}\right)$ are both minimum covers of a bipartite graph $K$ having finite cover number, then $B_{1}=B_{2}$.

The following theorem will be useful in identifying spent outputs in CryptoNote transaction graphs. We give a proof as it was not explicitly stated by Dulmage and Mendelsohn 5, although it follows from their results.

Theorem 5. Let $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ be minimum covers of a bipartite graph $K$ having finite cover number such that $A_{1} \subseteq A_{2}$. Then every maximum matching $M$ on $K$ has $\left|A_{2}\right|-\left|A_{1}\right|$ edges in the set $\left(A_{2} \backslash A_{1}\right) \times\left(B_{1} \backslash B_{2}\right)$.

Proof. Since $A_{1} \subseteq A_{2}$, Theorem 4 tells us that $B_{2} \subseteq B_{1}$. Since $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ are both minimum covers, Corollary 1 tells us that every maximum matching $M$ of $K$ has exactly $\left|A_{1}\right|$ edges in $A_{1} \times B_{1}^{c}$ and exactly $\left|A_{2}\right|$ edges in $A_{2} \times B_{2}^{c}$.

As $A_{1} \times B_{1}^{c} \subseteq A_{2} \times B_{2}^{c}$, every edge of $M$ in $A_{1} \times B_{1}^{c}$ is contained in $A_{2} \times B_{2}^{c}$. Thus $M$ has $\left|A_{2}\right|-\left|A_{1}\right|$ edges in $\left(A_{2} \times B_{2}^{c}\right) \backslash\left(A_{1} \times B_{1}^{c}\right)$. As illustrated in Fig. 5 , the set $A_{2} \times B_{2}^{c}$ can be partitioned as

$$
\begin{align*}
A_{2} \times B_{2}^{c}= & {\left[A_{1} \cup\left(A_{2} \backslash A_{1}\right)\right] \times\left[\left(B_{1} \cup B_{1}^{c}\right) \cap B_{2}^{c}\right] } \\
= & {\left[A_{1} \cup\left(A_{2} \backslash A_{1}\right)\right] \times\left[\left(B_{1} \backslash B_{2}\right) \cup B_{1}^{c}\right] } \\
= & {\left[A_{1} \times\left(B_{1} \backslash B_{2}\right)\right] \cup\left[A_{1} \times B_{1}^{c}\right] } \\
& \cup\left[\left(A_{2} \backslash A_{1}\right) \times\left(B_{1} \backslash B_{2}\right)\right] \cup\left[\left(A_{2} \backslash A_{1}\right) \times B_{1}^{c}\right] . \tag{3}
\end{align*}
$$

Since $\left(A_{1}, B_{1}\right)$ is a minimum cover and $A_{1} \times\left(B_{1} \backslash B_{2}\right) \subseteq A_{1} \times B_{1}$, Theorem 2 tells us that the matching $M$ cannot have any edges in $A_{1} \times\left(B_{1} \backslash B_{2}\right)$.

Since $\left(A_{1}, B_{1}\right)$ is a vertex cover and $\left(A_{2} \backslash A_{1}\right) \times B_{1}^{c}=\left(A_{2} \cap A_{1}^{c}\right) \times B_{1}^{c} \subseteq A_{1}^{c} \cap$ $B_{1}^{c}$, Lemma 1 tells us that the graph $K$ cannot have any edges in $\left(A_{2} \backslash A_{1}\right) \times B_{1}^{c}$. Consequently, the matching $M$ cannot have any edges in this set.

The above observations tell us that two of the partition elements in equation (3) cannot have edges from a maximum matching $M$. Thus the $\left|A_{2}\right|-\left|A_{1}\right|$ edges of $M$ in $\left(A_{2} \times B_{2}^{c}\right) \backslash\left(A_{1} \times B_{1}^{c}\right)$ must belong to $\left(A_{2} \backslash A_{1}\right) \times\left(B_{1} \backslash B_{2}\right)$.

If a matching on a graph has an edge incident on a vertex, we say that the vertex is matched by the matching. The following corollary of Theorem 5 says that all the vertices in the difference between two minimum covers are matched by every maximum matching.


Fig. 5. Partition of $A_{2} \times B_{2}^{c}$ in the proof of Theorem 5

Corollary 3. Let $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ be minimum covers of a bipartite graph $K$ having finite cover number such that $A_{1} \subseteq A_{2}$. Let $M$ be any maximum matching on $K$. Then all the vertices in the sets $A_{2} \backslash A_{1}$ and $B_{1} \backslash B_{2}$ are matched by $M$.

Proof. Theorem 5 tells us that any maximum matching $M$ must have $\left|A_{2}\right|-\left|A_{1}\right|$ edges in the set $\left(A_{2} \backslash A_{1}\right) \times\left(B_{1} \backslash B_{2}\right)$. Since $A_{1} \subseteq A_{2},\left|A_{2} \backslash A_{1}\right|=\left|A_{2}\right|-\left|A_{1}\right|$. As any two distinct edges in $M$ cannot have a vertex in common, each vertex in $A_{2} \backslash A_{1}$ must have exactly one of the $\left|A_{2}\right|-\left|A_{1}\right|$ edges of $M$ incident on it.

Since $\left(A_{1}, B_{1}\right)$ and $\left(A_{2}, B_{2}\right)$ are both minimum covers, $\left|A_{1}\right|+\left|B_{1}\right|=\left|A_{2}\right|+$ $\left|B_{2}\right| \Longrightarrow\left|B_{1}\right|-\left|B_{2}\right|=\left|A_{2}\right|-\left|A_{1}\right|$. As $B_{1} \subseteq B_{2},\left|B_{1} \backslash B_{2}\right|=\left|B_{1}\right|-\left|B_{2}\right|$. Thus each vertex in $B_{1} \backslash B_{2}$ has exactly one of the $\left|A_{2}\right|-\left|A_{1}\right|$ edges of $M$ incident on it.

By this corollary, if we can find two distinct minimum covers of the transaction graph induced by a CryptoNote blockchain, then we would have identified some outputs which are matched by every maximum matching on this graph. Thus every candidate for the true association between outputs and key images has these outputs marked as spent.

For a bipartite graph $K$, let $\mathcal{C}$ be the set of all minimum covers. Let us define the following sets obtained by taking intersections and unions of the components
of the minimum covers.

$$
\begin{align*}
& A_{*}=\bigcap_{(A, B) \in \mathcal{C}} A, \quad A^{*}=\bigcup_{(A, B) \in \mathcal{C}} A  \tag{4}\\
& B_{*}=\bigcap_{(A, B) \in \mathcal{C}} B, \quad B^{*}=\bigcup_{(A, B) \in \mathcal{C}} B \tag{5}
\end{align*}
$$

By Theorem 3, if $K$ has a finite cover number then the pairs $\left(A_{*}, B^{*}\right)$ and $\left(A^{*}, B_{*}\right)$ are both minimum covers of $K$.

Example 2. Consider the bipartite graph in Fig. 3 with vertex classes $S=$ $\left\{P_{1}, P_{2}, P_{3}, P_{4}\right\}$ and $T=\left\{I_{1}, I_{2}, I_{3}\right\}$. Since $(\emptyset, T)$ is a minimum cover the graph, $A_{*}=\emptyset$ and $B^{*}=T$. As $\left(\left\{P_{1}\right\},\left\{I_{2}, I_{3}\right\}\right)$ and $\left(\left\{P_{1}, P_{2}\right\},\left\{I_{3}\right\}\right)$ are the only other minimum covers, $A^{*}=\left\{P_{1}, P_{2}\right\}$ and $B_{*}=\left\{I_{3}\right\}$.

With the above definitions in place, we are ready to describe the DulmageMendelsohn (DM) decomposition.

Definition 9. Let $K=(S, T, E)$ be a bipartite graph having a finite cover number. The Dulmage-Mendelsohn decomposition of $K$ is a partition of $S \times T$ into three disjoint sets $R_{1}, R_{2}, R_{3}$ which satisfy the following properties:

1. The set of admissible edges in $K$ equals $E \cap R_{1}$.
2. The set of inadmissible edges in $K$ equals $E \cap R_{2}$.
3. $E \cap R_{3}=\emptyset$.

The structure of the sets $R_{1}, R_{2}, R_{3}$ depends on the minimum covers of $K$. Suppose $A_{*}=A^{*}$. Then the graph $K$ has only one minimum cover given by $\left(A_{*}, B^{*}\right)=\left(A^{*}, B_{*}\right)$. In this case, the sets $R_{1}, R_{2}, R_{3}$ are given by

$$
\begin{align*}
& R_{1}=\left(A_{*} \times\left(B^{*}\right)^{c}\right) \bigcup\left(\left(A_{*}\right)^{c} \times B^{*}\right) \\
& R_{2}=A_{*} \times B^{*}  \tag{6}\\
& R_{3}=\left(A_{*}\right)^{c} \times\left(B^{*}\right)^{c}
\end{align*}
$$

It is clear that $S \times T=R_{1} \cup R_{2} \cup R_{3}$ and $R_{i} \cap R_{j}=\emptyset$ for $i \neq j, 1 \leq i, j \leq 3$. Since $\left(A_{*}, B^{*}\right)$ is a vertex cover, Lemma 1 tells us that $E \cap R_{3}=\emptyset$. As $\left(A_{*}, B^{*}\right)$ is the only possible minimum cover of $K$, Theorem 2 tells us that the set of inadmissible edges in $K$ equals $E \cap R_{2}$. Furthermore, the theorem tells us that the edges in $E \cap R_{2}^{c}$ are admissible. Since there are no edges in $E \cap R_{3}$, the set of admissible edges in $K$ equals $E \cap R_{1}$.

Now suppose $A_{*} \neq A^{*}$. By definition, $A_{*} \subseteq A^{*}$. So $A_{*}$ must be a proper subset of $A^{*}$. Then there exists at least one non-empty set $X \subset S$ such that $A_{*} \cap X=\emptyset$ and $\left(A_{*} \cup X, Y\right)$ is a minimum cover of $K$ for some $Y \subset T$. The existence of such a set follows from the fact $A^{*} \backslash A_{*}$ is a candidate for $X$. Let $S_{1}$ be a set of smallest cardinality among all candidates for $X$. There may be many possibilities for $S_{1}$, all having the same smallest cardinality. We can pick any one of them.

Let $\left(A_{1}, B_{1}\right)$ be a minimum cover with $A_{1}=A_{*} \cup S_{1}$. By Corollary 2, $B_{1}$ is uniquely determined by $A_{1}$. As $A_{*} \subseteq A_{1}$, Theorem 4 tells us that $B_{1} \subseteq B^{*}$. As all minimum covers of $K$ have the same size, we have $\left|A_{*}\right|+\left|B^{*}\right|=\left|A_{1}\right|+\left|B_{1}\right|$. Since $\left|A_{1}\right|>\left|A_{*}\right|$, we have $\left|B_{1}\right|<\left|B^{*}\right|$. Thus $B_{1}$ is a proper subset of $B^{*}$. Let $T_{1}=B^{*} \backslash B_{1}$. Since $\left|A_{1}\right|-\left|A_{*}\right|=\left|B^{*}\right|-\left|B_{1}\right|$, we have $\left|S_{1}\right|=\left|T_{1}\right|$.

If $A_{1}=A^{*}$, the process stops. Otherwise, there exists at least one nonempty set $X \subset S$ such that $A_{1} \cap X=\emptyset$ and $A_{1} \cup X$ is the first component of a minimum cover of $K$. Let $S_{2}$ be a set of smallest cardinality among all candidates for $X$. Let $\left(A_{2}, B_{2}\right)$ be a minimum cover with $A_{2}=A_{1} \cup S_{2}=A_{*} \cup S_{1} \cup S_{2}$. As before, $B_{2}$ is uniquely determined by $A_{2}$ and $B_{2} \subset B_{1}$. Let $T_{2}=B_{1} \backslash B_{2}$. Since $\left|A_{2}\right|-\left|A_{1}\right|=\left|B_{1}\right|-\left|B_{2}\right|$, we have $\left|S_{2}\right|=\left|T_{2}\right|$. Since $B^{*}=T_{1} \cup B_{1}$ and $T_{2}=B_{1} \backslash B_{2}$, we have $B^{*}=T_{1} \cup T_{2} \cup B_{2}$.

If we proceed in this manner, the process will stop for some $k$ where

$$
\begin{equation*}
A_{*} \cup S_{1} \cup S_{2} \ldots \cup S_{k}=A^{*} \tag{7}
\end{equation*}
$$

At this point, $\left(A^{*}, B_{*}\right)$ will be the resulting minimum cover. Furthermore, the $T_{i}$ 's satisfy

$$
\begin{equation*}
B^{*}=T_{1} \cup T_{2} \cup \ldots T_{k} \cup B_{*} \tag{8}
\end{equation*}
$$

In the intermediate stages of this process, $\left(A_{i}, B_{i}\right)$ is a minimum cover for $K$ for each $i \in\{1,2, \ldots, k\}$ where

$$
\begin{align*}
& A_{i}=A_{*} \cup S_{1} \cup S_{2} \cup \ldots \cup S_{i},  \tag{9}\\
& B_{i}=T_{i+1} \cup T_{i+2} \cup \ldots \cup T_{k} \cup B_{*} . \tag{10}
\end{align*}
$$

Equations (7) and (8) give the following decompositions of the vertex classes $S$ and $T$.

$$
\begin{align*}
& S=A^{*} \bigcup\left(A^{*}\right)^{c}=A_{*} \cup S_{1} \cup S_{2} \ldots \cup S_{k} \bigcup\left(A^{*}\right)^{c}  \tag{11}\\
& T=\left(B^{*}\right)^{c} \bigcup B^{*}=\left(B^{*}\right)^{c} \bigcup T_{1} \cup T_{2} \ldots \cup T_{k} \cup B_{*} . \tag{12}
\end{align*}
$$

The $k+2$ sets in the unions on the extreme right of both the above equations form a partition of $S$ and $T$ respectively. These partitions are unique except for a permutation of the $S_{i}$ 's having same cardinality, with the $T_{i}$ 's appropriately permuted.

We claim that the DM decomposition is given by the sets

$$
\begin{align*}
& R_{1}=\left(A_{*} \times\left(B^{*}\right)^{c}\right) \bigcup\left(S_{1} \times T_{1}\right) \bigcup \ldots \bigcup\left(S_{k} \times T_{k}\right) \bigcup\left(\left(A^{*}\right)^{c} \times B_{*}\right)  \tag{13}\\
& R_{2}=\left(A_{*} \times B^{*}\right) \bigcup\left(A^{*} \times B_{*}\right) \bigcup_{i<j}\left(S_{i} \times T_{j}\right)  \tag{14}\\
& R_{3}=\left(\left(A_{*}\right)^{c} \times\left(B^{*}\right)^{c}\right) \bigcup\left(\left(A^{*}\right)^{c} \times\left(B_{*}\right)^{c}\right) \bigcup_{i>j}\left(S_{i} \times T_{j}\right) \tag{15}
\end{align*}
$$

To visualize the DM decomposition, suppose we order the vertices in $S$ according to the partition in equation (11), i.e. the vertices in $A_{*}$ appear first, followed


Fig. 6. The DM decomposition of a graph.
by vertices in $S_{1}, S_{2}, \ldots, S_{k}$, and $\left(A^{*}\right)^{c}$. Similarly, suppose the vertices in $T$ are ordered according to the partition in equation 12 . Then the DM decomposition can be represented by Fig. 6, where the rows correspond to vertices in $T$ and the columns correspond to vertices in $S$. The admissible edges lie in blocks along the diagonal, the inadmissible edges lie above these blocks, and there are no edges below these blocks.

Furthermore, by applying Theorem 5 to adjacent minimum covers in the sequence $\left(A_{*}, B^{*}\right),\left(A_{1}, B_{1}\right), \ldots,\left(A_{k-1}, B_{k-1}\right),\left(A^{*}, B_{*}\right)$, we conclude that every maximum matching on the graph has $\left|S_{i}\right|$ edges in $S_{i} \times T_{i}$ for $i=1,2, \ldots, k$. By Corollary 3, every maximum matching matches all the vertices in $S_{i}$ and $T_{i}$. If we choose $S$ to be the set of all outputs in the bipartite graph induced by a CryptoNote blockchain transaction history, then the sets $S_{i}$ contain only spent outputs for all $i=1,2, \ldots, k$.

Before checking that the sets $R_{1}, R_{2}, R_{3}$ satisfy the properties required of a DM decomposition, let us calculate them for our running example.

Example 3. Consider the bipartite graph in Fig. 3 with vertex classes $S=$ $\left\{P_{1}, P_{2}, P_{3}, P_{4}\right\}$ and $T=\left\{I_{1}, I_{2}, I_{3}\right\}$.

- As we noted in Example 2, $A_{*}=\emptyset, B^{*}=T$ and $A^{*}=\left\{P_{1}, P_{2}\right\}, B_{*}=\left\{I_{3}\right\}$.
- As $\left(\left\{P_{1}\right\},\left\{I_{2}, I_{3}\right\}\right)$ is the only candidate for $\left(A_{1}, B_{1}\right)$, we have $S_{1}=\left\{P_{1}\right\}$ and $T_{1}=\left\{I_{1}\right\}$.
$-\operatorname{As}\left(\left\{P_{1}, P_{2}\right\},\left\{I_{3}\right\}\right)$ is the only candidate for $\left(A_{2}, B_{2}\right)$, we have $S_{2}=\left\{P_{2}\right\}$ and $T_{2}=\left\{I_{2}\right\}$.

The DM decomposition is given by

$$
\begin{aligned}
& R_{1}=\left\{\left(P_{1}, I_{1}\right),\left(P_{2}, I_{2}\right),\left(P_{3}, I_{3}\right),\left(P_{4}, I_{3}\right)\right\} \\
& R_{2}=\left\{\left(P_{1}, I_{3}\right),\left(P_{2}, I_{3}\right),\left(P_{1}, I_{2}\right)\right\} \\
& R_{3}=\left\{\left(P_{3}, I_{1}\right),\left(P_{3}, I_{2}\right),\left(P_{4}, I_{1}\right),\left(P_{4}, I_{2}\right),\left(P_{2}, I_{1}\right)\right\}
\end{aligned}
$$

The graph has no edges in $R_{3}$. The edges in $R_{2}$ cannot appear in any maximum matching, as $P_{1}$ must be matched to $I_{1}$ and $P_{2}$ must be matched to $I_{2}$. The edges in $R_{1}$ appear in at least one maximum matching on the graph. Every maximum matching has one edge in $S_{1} \times T_{1}=\left\{\left(P_{1}, I_{1}\right)\right\}$ and one edge in $S_{2} \times T_{2}=$ $\left\{\left(P_{2}, I_{2}\right)\right\}$. Finally, $S_{1}$ and $S_{2}$ contain only spent outputs.

Returning to the general DM decomposition given in equations (13), (14), (15), let us first show that the graph cannot have edges in the set $R_{3}$. Since $\left(A_{*}, B^{*}\right)$ and $\left(A^{*}, B_{*}\right)$ are vertex covers, Lemma 1 tells us that the graph has no edges in $\left(A_{*}\right)^{c} \times\left(B^{*}\right)^{c}$ and $\left(A^{*}\right)^{c} \times\left(B_{*}\right)^{c}$. For $i \geq 2$, we have

$$
\begin{align*}
& A_{i-1}^{c}=S \backslash A_{i-1}=S_{i} \cup S_{i+1} \ldots \cup S_{k} \bigcup\left(A^{*}\right)^{c},  \tag{16}\\
& B_{i-1}^{c}=T \backslash B_{i-1}=\left(B^{*}\right)^{c} \bigcup T_{1} \cup T_{2} \ldots \cup T_{i-1}, \tag{17}
\end{align*}
$$

as seen by the representations of $A_{i}, B_{i}$ in equations (9), 10 and the representations of $S, T$ in equations (11), 12 . For $i>j$, each $S_{i} \times T_{j}$ is contained in $\left(A_{i-1}^{c}, B_{i-1}^{c}\right)$. As $\left(A_{i-1}, B_{i-1}\right)$ is vertex cover, by Lemma 1 the graph cannot have edges in $S_{i} \times T_{j}$ for $i>j$. This completes the proof that the edge set $E$ of the graph $K$ satisfies $E \cap R_{3}=\emptyset$ for the $R_{3}$ in equation 15 .

Now let us show that the set of inadmissible edges in $K$ equals $E \cap R_{2}$ for the $R_{2}$ given in equation 14 . Since $\left(A_{*}, B^{*}\right)$ and $\left(A^{*}, B_{*}\right)$ are minimum covers of the graph, Theorem 2 tells us that graph edges in $A_{*} \times B^{*}$ and $A^{*} \times B_{*}$ are inadmissible. Observe that for $i<j$ the set $S_{i} \times T_{j}$ is contained in $A_{i} \times B_{i}$, as seen in equations (9), 10). As each $\left(A_{i}, B_{i}\right)$ is a minimum cover of the graph, Theorem 2 once again tells us that graph edges in $S_{i} \times T_{j}$ for $i<j$ are inadmissible. But these results merely tell us that $E \cap R_{2}$ is a subset of the set of inadmissible edges. We want to show that it equals the set of inadmissible edges in $K$.

In the case of $A_{*}=A^{*}$, the graph had only one minimum cover $\left(A_{*}, B^{*}\right)$, which simplified the task of finding the set of inadmissible edges. For $A_{*} \neq A^{*}$, there could be minimum covers $(A, B)$ which are not equal to any of $\left(A_{*}, B^{*}\right)$, $\left(A_{1}, B_{1}\right),\left(A_{2}, B_{2}\right), \ldots,\left(A_{k-1}, B_{k-1}\right),\left(A^{*}, B_{*}\right)$. However, in their 1958 paper [5], Dulmage and Mendelsohn proved that the any minimum cover $(A, B)$ of $K$ can be represented by a combination of the $S_{i}$ 's and $T_{i}$ 's as described in the following theorem.

Theorem 6. For a bipartite graph $K$ having a finite cover number, let $A_{*}, B_{*}$, $S_{1}, S_{2}, \ldots, S_{k}, T_{1}, T_{2}, \ldots, T_{k}$ be the sets obtained in the procedure described earlier in this section. Let $(A, B)$ be any minimum cover of $K$. Then there exist complementary subsets $\Delta$ and $\Pi$ of $\{1,2, \ldots, k\}$ such that

$$
\begin{aligned}
& A=A_{*} \bigcup\left(\bigcup_{i \in \Delta} S_{i}\right) \\
& B=\left(\bigcup_{j \in \Pi} T_{j}\right) \bigcup B_{*}
\end{aligned}
$$

This theorem (in combination with Theorem 2) tells us that the set of inadmissible edges equals the union of $E \cap(A \times B)$ as the set $\Delta$ varies over the $2^{k}$ subsets of $\{1,2, \ldots, k\}$ with $\Pi=\Delta^{c}$. But all such sets $E \cap(A \times B)$ are contained in $R_{2}$. To see this, note that

$$
\begin{align*}
E \cap[A \times B] & =E \cap\left[\left(A_{*} \times B\right) \bigcup\left(\cup_{i \in \Delta} S_{i} \times B\right)\right] \\
& =E \cap\left[\left(A_{*} \times B\right) \bigcup\left(\cup_{i \in \Delta} S_{i} \times B_{*}\right) \bigcup\left(\cup_{i \in \Delta} S_{i} \times \cup_{j \in \Pi} T_{j}\right)\right] \\
& \subseteq E \cap\left[\left(A_{*} \times B^{*}\right) \bigcup\left(A^{*} \times B_{*}\right) \bigcup\left(\cup_{i \in \Delta} S_{i} \times \cup_{j \in \Pi} T_{j}\right)\right]  \tag{18}\\
& =E \cap\left[\left(A_{*} \times B^{*}\right) \bigcup\left(A^{*} \times B_{*}\right) \bigcup_{i \in \Delta, j \in \Pi, i<j}\left(S_{i} \times T_{j}\right)\right]  \tag{19}\\
& \subseteq E \cap R_{2}, \tag{20}
\end{align*}
$$

where the subset relation in $(18)$ follows from equations $(7)$ and $(8)$ which show that $\cup_{i \in \Delta} S_{i} \subseteq A^{*}$ and $B \subseteq B^{*}$. The equality in 19 follows from two observations: (a) the graph cannot have edges in $S_{i} \times T_{j}$ for $i>j$, as discussed in our argument showing $E \cap R_{3}=\emptyset$, and (b) $i \neq j$ when $i \in \Delta$ and $j \in \Pi=\Delta^{c}$. The subset relation in (20) follows from definition of $R_{2}$ in 14 . Thus, we conclude that the set of inadmissible edges in the graph $K$ equals $E \cap R_{2}$.

Finally, as $R_{1}, R_{2}, R_{3}$ form a partition of $S \times T$ with $E \cap R_{3}=\emptyset$, the set of admissible edges must equal $E \cap R_{2}^{c}$ which is equal to $E \cap R_{1}$. This completes the proof that the expressions for $R_{1}, R_{2}, R_{3}$ in equations (13), 14, (15), satisfy the properties of a DM decomposition given in Definition 9 .

## Computing the DM Decomposition

The DM decomposition of a bipartite graph $K$ can be computed by finding a maximum matching $M$ on $K$, then finding subsets of vertex classes unreachable from $M$ via alternating paths, and finally by finding strongly connected components of the subgraph induced by the unreachable vertices (see [12] for details). Both open source [4] and proprietary [1] implementations of the DM decomposition algorithm are available.

## 3 DM Decomposition of the Monero Graph

A transaction ring is said to be traceable if the true output being spent is identified. To evaluate the effectiveness of the DM decomposition in tracing transaction rings, we used the results obtained by Yu et al. [16] on Monero as the benchmark. The latter results are the best results on Monero traceability which do not use information from hard forks like Monero Original and MoneroV. Hinteregger et al. 7] used the key images which appeared in both the main Monero chain and these hard forks to trace transactions in all three chains. We were unable to download the MoneroV and Monero Original blockchain data. This prevented us from using the cross-chain data in evaluating the DM decomposition 1

The algorithm proposed by Yu et al. [16] first applies the cascade algorithm proposed by Kumar et al. [8] and Moser et al. [10]. Then it uses a clustering algorithm to find sets of spent outputs, called closed sets (these correspond to the $S_{i}$ 's in the DM decomposition). The algorithm classifies a transaction ring as traceable if the cascade/clustering algorithms mark all outputs in the ring except one as spent in another ring. Similarly, the DM decomposition classifies a transaction ring as traceable if only one edge incident on the corresponding key image is admissible.

Yu et al. considered Monero transactions contained in blocks with height up to $1,541,236$ (March 30, 2018). This data set contains $23,164,745$ transaction rings (each one contributing a key image) and $25,126,033$ outputs. The corresponding bipartite graph has $58,791,856$ edges. In Monero, RingCT outputs have amounts hidden in Pedersen commitments. They were introduced in Monero in January 2017 and became mandatory in September 2017 [3]. Out of the $23,164,175$ transaction rings in the data set, $4,330,234$ were RingCT rings and the remaining $18,834,511$ were pre-RingCT rings.

Previous work [8], [10], [16], on Monero traceability has shown that RingCT transactions in Monero are immune to traceability attacks. The same observation holds for the DM decomposition approach. None of the 4,330,234 RingCT rings could be traced by the DM decomposition. Table 1 compares the number of preRingCT transaction rings traced by the cascade/clustering (CC) algorithm and the DM decomposition. Each row in the table gives results for transaction rings which have a certain number of mixin outputs. The results for all transaction rings with 10 or more mixin outputs are combined in the row with label " $\geq 10$ ".

All the $16,335,308$ rings traced by the DM decomposition are associated with a set $S_{i}$ with $\left|S_{i}\right|=1$. The singleton set $T_{i}$ corresponding to $S_{i}$ has the key image of the output in $S_{i}$. As seen from the last row, the DM decomposition identifies only 341 more traceable rings than the CC algorithm. These new rings are only among the transaction rings having 2,3 , or 4 mixins. Thus, for transactions up to block height $1,541,236$ the advantage of using the DM decomposition for tracing Monero transactions is marginal.

[^0]| No. of mixins | No. of rings | Traced by CC | Traced by DM |
| :---: | ---: | ---: | ---: |
| 0 | 12209675 | 12209675 | 12209675 |
| 1 | 707786 | 625641 | 625641 |
| 2 | 4496490 | 1779134 | 1779446 |
| 3 | 1486593 | 952855 | 952862 |
| 4 | 3242625 | 451959 | 451981 |
| 5 | 319352 | 74186 | 74186 |
| 6 | 432875 | 202360 | 202360 |
| 7 | 21528 | 4296 | 4296 |
| 8 | 30067 | 3506 | 3506 |
| 9 | 17724 | 2178 | 2178 |
| $\geq 10$ | 200030 | 29177 | 29177 |
| Total | 23164745 | 16334967 | 16335308 |

Table 1. Monero traceability of pre-RingCT rings by the CC algorithm vs DM decomposition (up to block $1,541,236$ )

Yu et al. report finding 3017 closed sets with sizes in the range 2 to 55 . In the DM decomposition, each $S_{i}$ is a closed set. The DM decomposition is able to find 3045 closed sets with 3041 of them having sizes in the range 2 to 55 . The remaining four closed sets have sizes $103,106,119$, and 122 .

The DM decomposition marked $15,633,140$ out of the $58,791,856$ edges in the bipartite graph as inadmissible. Each inadmissible edge reduces the effective mixin size of a transaction ring. Table 2 gives the counts of transactions with a certain number of mixins before and after the DM decomposition $2^{2}$ As expected, transaction rings with smaller effective mixin size are more frequent after the DM decomposition.

To check if the transactions which have appeared after block 1,541,236 have affected the traceability of RingCT rings, we computed the DM decomposition of the subgraph induced exclusively by RingCT transaction rings in all blocks up to height 2,330,000 (April 1, 2021). This subgraph has 26,098,794 key images and $29,588,617$ outputs with $252,843,948$ edges between them. Let $\mathcal{K}$ be the set of all the key images in this subgraph. Its DM decomposition revealed only two minimum covers, $(\emptyset, \mathcal{K})$ and $\left(S_{1}, \mathcal{K} \backslash T_{1}\right)$ where $\left|S_{1}\right|=\left|T_{1}\right|=5$. The set $S_{1}$ consists of RingCT outputs with indices $3890287,3890288,3890289$, 3890290 , and 3890291. These five outputs were created by Wijaya et al. [14] in block $1,468,425]^{3}$ All of them were spent using the other four as mixins in five transaction rings in block $1,468,439$ (Dec 17, 2017), to demonstrate that a set of outputs can be considered spent without relying on zero-mixin transactions ${ }^{4}$. These five outputs are also marked as spent by the Monero blackball tool [11]. Thus, the DM decomposition

[^1]| Effective <br> no. of mixins | No. of rings <br> before DMD | No. of rings <br> after DMD |
| :---: | ---: | ---: |
| 0 | 12209675 | 16335308 |
| 1 | 707786 | 1413028 |
| 2 | 4496490 | 2369796 |
| 3 | 1486593 | 279377 |
| 4 | 3242625 | 2369578 |
| 5 | 319352 | 186257 |
| 6 | 432875 | 73690 |
| 7 | 21528 | 13086 |
| 8 | 30067 | 23615 |
| 9 | 17724 | 13071 |
| $\geq 10$ | 200030 | 87939 |
| Total | 23164745 | 23164745 |

Table 2. Effective number of mixins before and after DM decomposition (up to block $1,541,236)$. Only 17 RingCT rings experience a change in effective number of mixins.
of the Monero RingCT subgraph (using only main chain data) does not identify any new outputs as spent.

There were $22,785,298$ RingCT transaction rings in the blocks with heights from $1,468,426$ to $2,330,000$. The five spent RingCT outputs were chosen as mixins in only 17 of these RingCT rings. The block heights and RingCT ring indices of the affected rings are shown in Table 3 Each of the 17 rings has its effective number of mixins reduced by one. The latest affected ring appears in block 1,521,556 (March 3, 2018). Thus, the change in effective number of mixins shown in Table 2 is mostly in pre-RingCT rings.

Justin Ehrenhofer maintains a list of RingCT outputs which are known to be spent [2]. This list was generated from both hard fork data and mining payouts. It contains the spent outputs identified by Hinteregger et al. as a subset ${ }^{5}$ It is meant to be consumed by Monero wallets to avoid picking these spent outputs as mixins. The data set does not mention the transaction rings in which each output was spent. If this information were available, some edges from the RingCT subgraph can be removed, leading to a different DM decomposition. We hope to obtain this information from the Monero community and share our findings at a later time.

## 4 Conclusion

We have described how the Dulmage-Mendelsohn decomposition of bipartite graphs can be used to characterize the information revealed by CryptoNote transaction rings. It is surprising that this decomposition has gone unnoticed for so long, as the idea of maximum matchings on CryptoNote transaction graphs

[^2]| Block <br> height | RingCT output indices in affected ring | Mixin output <br> index |
| :---: | :--- | :---: |
| 1468459 | $2598830,3003977,3355066,3434937,3890288$ | 3890288 |
| 1468463 | $2547881,3767909,3872300,3882612,3890290$ | 3890290 |
| 1468528 | $2547174,3038956,3635398,3806854,3890287$ | 3890287 |
| 1468554 | $1994254,3214788,3682735,3870916,3890291$ | 3890291 |
| 1468589 | $2293356,2959042,3502654,3853789,3890290$ | 3890290 |
| 1468589 | $1275085,3315022,3439832,3748016,3890289$ | 3890289 |
| 1468610 | $3846244,3873104,3875091,3881368,3890291,3892572$ | 3890291 |
| 1468633 | $603380,3882040,3888806,3890287,3896742$ | 3890287 |
| 1468736 | $3194098,3862759,3890287,3895198,3896533$ | 3890287 |
| 1469040 | $3861157,3890289,3890864,3894739,3910002$ | 3890289 |
| 1469132 | $3452429,3875789,3884238,3890289,3910975$ | 3890289 |
| 1469171 | $2145284,2683687,3890287,3893435,3913073$ | 3890287 |
| 1491145 | $112916,1691643,2371432,3890287,3952071$ | 3890287 |
| 1497524 | $1463611,3890288,4205669,4265964,4292176,4487224,4491812,4496125,4496544$ | 3890288 |
| 1500511 | $3391172,3890290,4326731,4541461,4544110$ | 3890290 |
| 1508728 | $1464801,2033769,3890288,4243990,4533972,4654036,4657845,4660726,4663215$ | 3890288 |
| 1521556 | $735259,1038734,2887471,3890289,4791022,4842288$ | 3890289 |

Table 3. The 17 RingCT transaction rings affected by the 5 spent outputs (up to block $2,330,000$ )
being plausible candidates for the ground truth has been known for a while. The general form of the DM decomposition (the Gallai-Edmonds decomposition) is described as a central result of matching theory in the preface of the standard reference on the subject [9]. While the decomposition does not reveal much more about Monero than what was known before, it is preferable as it avoids the heuristics and computational bottlenecks of previous methods.

A natural question arises: How should the mixin sampling strategy in CryptoNote blockchains be designed to avoid revealing information via the DM decomposition? We do not have an answer. Empirically, the existing sampling strategy in Monero seems to be robust to the decomposition. Can one expect this robustness to continue in the future? Yu et al. [16] gave estimates on the probability of existence of a closed set for a uniform sampling strategy when each ring has 3 mixins. Similar analyses with more realistic assumptions are needed to understand the information leakage risks of the sampling strategies used in practice.

## References

1. dmperm: MATLAB function for Dulmage-Mendelsohn decomposition, https:// in.mathworks.com/help/matlab/ref/dmperm.html
2. Monero Blackball Databases, https://github.com/monero-blackball/ monero-blackball-site
3. Monero Scheduled Software Upgrades, https://github.com/monero-project/ monero/\#scheduled-software-upgrades
4. Davis, T.A.: CSparse: A concise sparse matrix package, https://people.engr. tamu.edu/davis/suitesparse.html
5. Dulmage, A.L., Mendelsohn, N.S.: Coverings of bipartite graphs. Canadian Journal of Mathematics 10, 517-534 (1958). https://doi.org/10.4153/CJM-1958-052-0
6. Goodell, B.: Perfect privacy or strong deniability? In: Monero Konferenco (2019), https://youtu.be/xicn4rdUj_Q
7. Hinteregger, A., Haslhofer, B.: An empirical analysis of Monero cross-chain traceability. In: Financial Cryptography and Data Security. pp. 150-157 (2019)
8. Kumar, A., Fischer, C., Tople, S., Saxena, P.: A traceability analysis of Monero's blockchain. In: European Symposium on Research in Computer Security. pp. 153173. Springer (2017)
9. Lovász, L., Plummer, M.: Matching Theory. North-Holland (1986)
10. Möser, M., Soska, K., Heilman, E., Lee, K., Heffan, H., Srivastava, S., Hogan, K., Hennessey, J., Miller, A., Narayanan, A., Christin, N.: An empirical analysis of traceability in the Monero blockchain. Proceedings on Privacy Enhancing Technologies 2018(3), 143-163 (2018). https://doi.org/10.1515/popets-2018-0025
11. Noether, S.: Sets of spent outputs. Monero Research Lab Technical Report MRL0007 (Nov 2018), https://www.getmonero.org/resources/research-lab/
12. Pothen, A., Fan, C.J.: Computing the block triangular form of a sparse matrix. ACM Trans. Math. Softw. 16(4), 303-324 (Dec 1990). https://doi.org/10.1145/98267.98287
13. Saberhagen, N.v.: CryptoNote v 2.0. White paper (2013), https://cryptonote. org/whitepaper.pdf
14. Wijaya, D.A., Liu, J., Steinfeld, R., Liu, D.: Monero ring attack: Recreating zero mixin transaction effect. In: 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). pp. 1196-1201 (2018). https://doi.org/10.1109/TrustCom/BigDataSE.2018.00165
15. Yu, J., Au, M.H.A., Esteves-Verissimo, P.: Re-thinking untraceability in the Cryptonote-style blockchain. In: 2019 IEEE 32nd Computer Security Foundations Symposium (CSF). pp. 94-106 (2019). https://doi.org/10.1109/CSF.2019.00014
16. Yu, Z., Au, M.H., Yu, J., Yang, R., Xu, Q., Lau, W.F.: New empirical traceability analysis of Cryptonote-style blockchains. In: Financial Cryptography and Data Security. pp. 133-149 (2019)

[^0]:    ${ }^{1}$ Hinteregger has made the public keys corresponding to spent outputs available on Zenodo https://zenodo.org/record/1304033 But we need the transaction rings where these outputs are spent to construct the graph.

[^1]:    ${ }^{2}$ Yu et al. presented the corresponding counts after execution of the CC algorithm in bar graph form. So we are unable to compare the exact numbers.
    $3 \begin{aligned} & 4 \\ & \text { https://xmrchain.net/tx/b6781f 2a6f5608553546442b84888346fdc3f78dd8995170180ed74081c05362 } \\ & \text { https://xmrchain.net/tx/8d4a0c7eccf92542eb5e1f09e72ccod934b180b768bc95388d33051db83194b }\end{aligned}$

[^2]:    5 https://github.com/oerpli/MONitERO/blob/master/csv/tx_spent.md

