Monero Cross-Chain Traceability

Empirical Analysis of Privacy Implications from Currency Hard-Forks

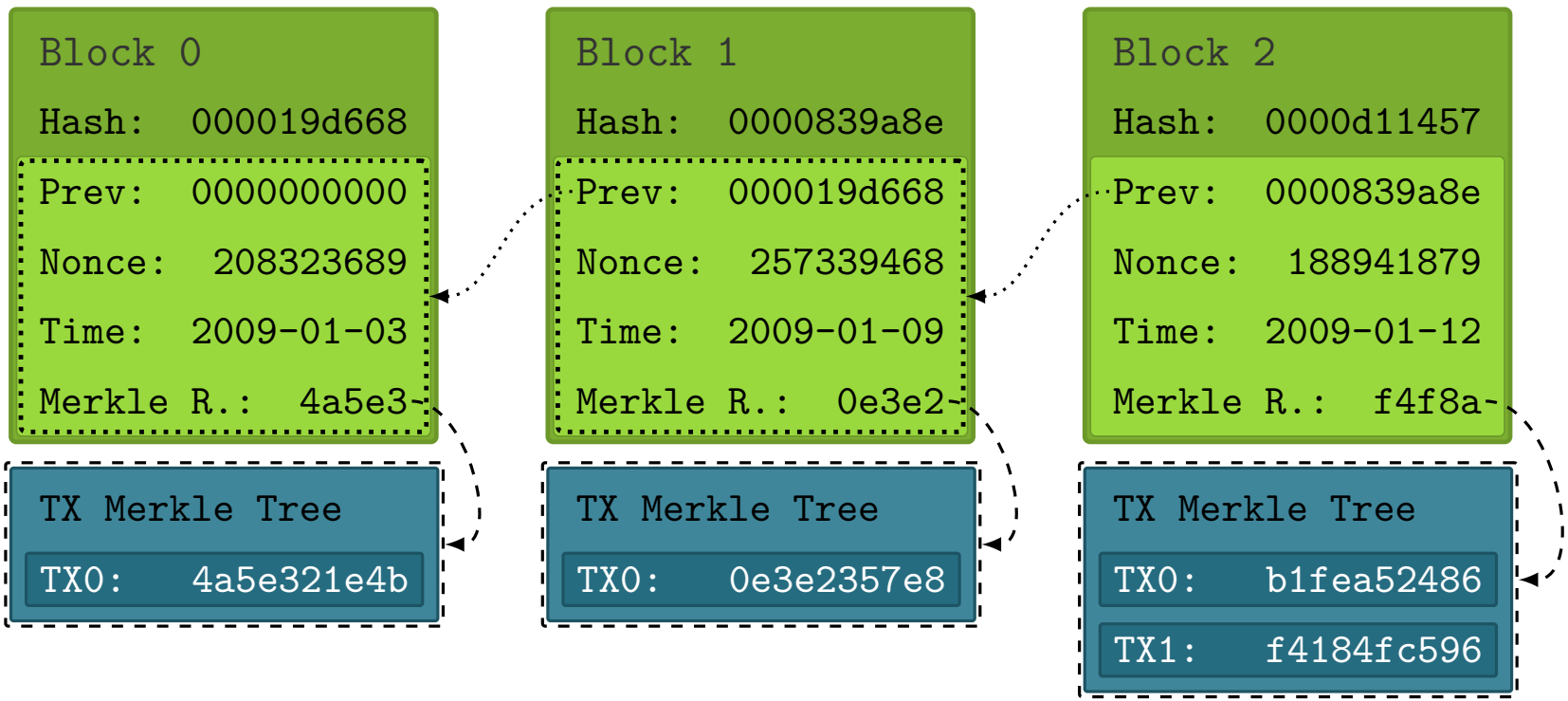
Abraham Hinteregger

November 13th, 2018

Cryptocurrencies

- [Nakamoto, 2008] laid the foundation of modern cryptocurrencies (scheme for trusted decentralized transactions)
- Transactions have inputs (references to previous outputs) and outputs
- Each output is issued to address of a user (public key)
- Recipient can spend outputs with private key

Blockchain


Block 0
Hash: 000019d668
Prev: 0000000000
Nonce: 208323689
Time: 2009-01-03
Merkle R.: 4a5e3-
TX Merkle Tree
TX0: 4a5e321e4b
Block 1
Hash: 0000839a8e
Prev: 000019d668
Nonce: 257339468
Time: 2009-01-09
Merkle R.: 0e3e2-
TX Merkle Tree
TX0: 0e3e2357e8
Block 2
Hash: 0000d11457
Prev: 0000839a8e
Nonce: 188941879
Time: 2009-01-12
Merkle R.: f4f8a-
TX Merkle Tree
TX0: b1fea52486
TX1: f4184fc596

Transaction

TX Hash: be83f7760b5f1a91

Version no: 1

#Inputs: 2

#Outputs: 2

TX Hash/Index: ba7521ec/2

Signature: 3045022100c...

TX Hash/Index: 888e0464/1

Signature: 30440220244...

Outputs:

0: Value: 1.99713455

Recipient addr: 126uLE1GDFxj

scriptPubKey: ...OP_CHECKSIG

1: Value: 6.00255800

Recipient addr: 16jaR3vF4TH3

scriptPubKey: ...OP_CHECKSIG

Monero

- Public nature of Bitcoin TX history prevents meaningful level of anonymity
- Monero (based on CryptoNote, [Van Saberhagen, 2013]) addresses this issue with the following methods:
  - Stealth Addresses (hide recipient addr.) $\rightarrow$ unlinkability (p. 26)

Monero

- Public nature of Bitcoin TX history prevents meaningful level of anonymity
- Monero (based on CryptoNote, [Van Saberhagen, 2013]) addresses this issue with the following methods:
  - Stealth Addresses (hide recipient addr.) $\rightarrow$ unlinkability (p. 26)
  - Ring Signatures (obfuscate spent outputs) $\rightarrow$ untraceability

Monero

- Public nature of Bitcoin TX history prevents meaningful level of anonymity
- Monero (based on CryptoNote, [Van Saberhagen, 2013]) addresses this issue with the following methods:
  - Stealth Addresses (hide recipient addr.) $\rightarrow$ unlinkability (p. 26)
  - Ring Signatures (obfuscate spent outputs) $\rightarrow$ untraceability
  - Confidential Transactions (hide amounts) $\rightarrow$ fungibility

Monero

- Public nature of Bitcoin TX history prevents meaningful level of anonymity
- Monero (based on CryptoNote, [Van Saberhagen, 2013]) addresses this issue with the following methods:
  - Stealth Addresses (hide recipient addr.) $\rightarrow$ unlinkability (p. 26)
  - Ring Signatures (obfuscate spent outputs) $\rightarrow$ untraceability
  - Confidential Transactions (hide amounts) $\rightarrow$ fungibility

Ring Signatures & Traceability

- Each TX input references:

Ring Signatures & Traceability

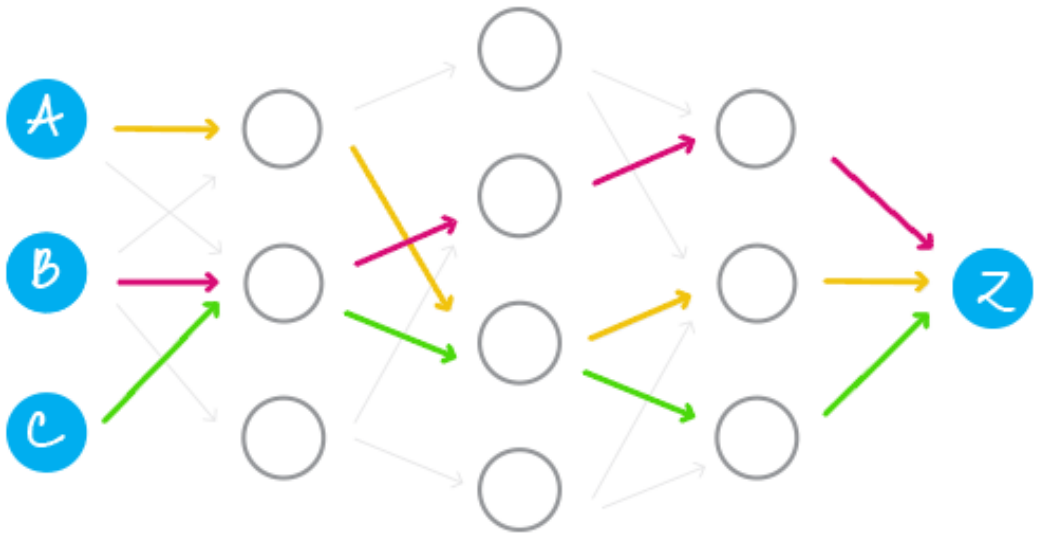
- Each TX input references:
  - Bitcoin: Output from older TX (TXO)

Ring Signatures & Traceability

- Each TX input references:
  - Bitcoin: Output from older TX (TXO)
  - Monero: Non-empty set of TXOs (a ring)

Ring Signatures & Traceability

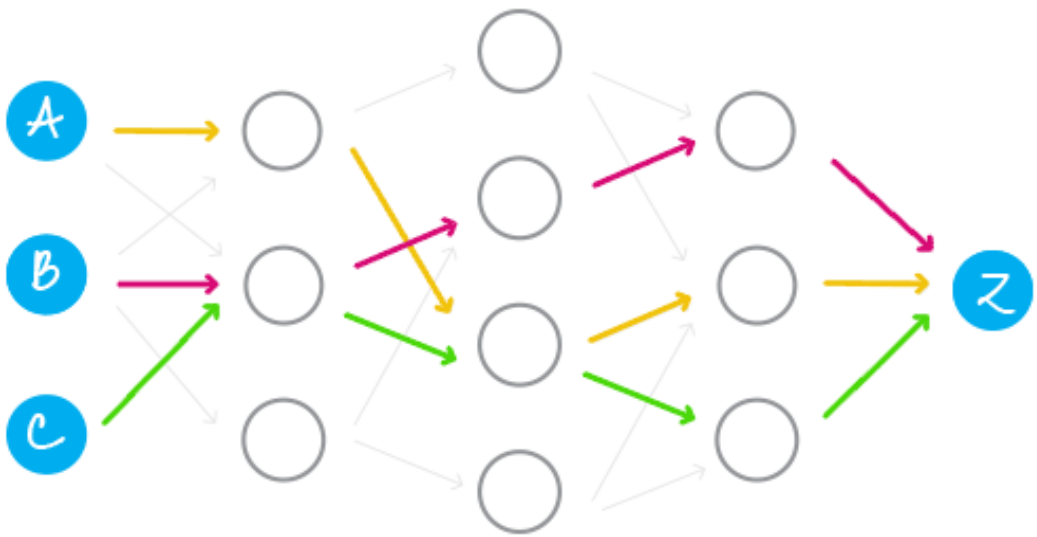
- Each TX input references:
  - Bitcoin: Output from older TX (TXO)
  - Monero: Non-empty set of TXOs (a ring)
- One ringmember is real, the others are decoys (mixins)



Source: `https://cryptonote.org/inside/`

Ring Signatures & Traceability

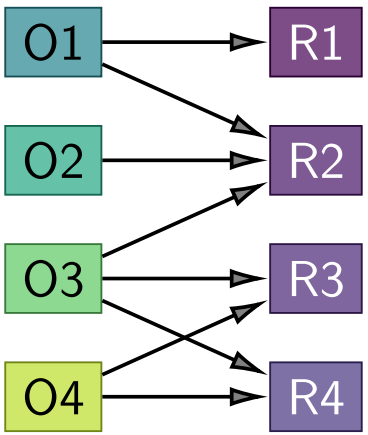
- Each TX input references:
  - Bitcoin: Output from older TX (TXO)
  - Monero: Non-empty set of TXOs (a ring)
- One ringmember is real, the others are decoys (mixins)



Source: `https://cryptonote.org/inside/`

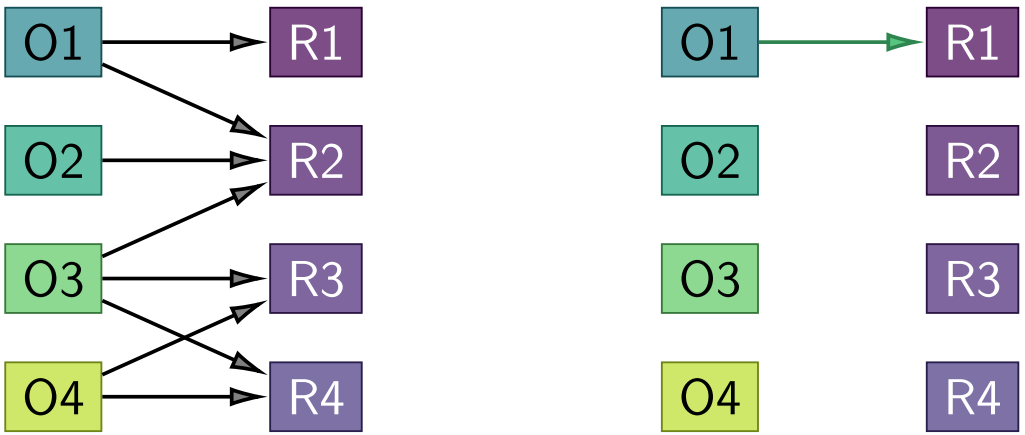
- Decoys are sampled from set of eligible outputs

Zero Mixin Removal & Intersection Removal



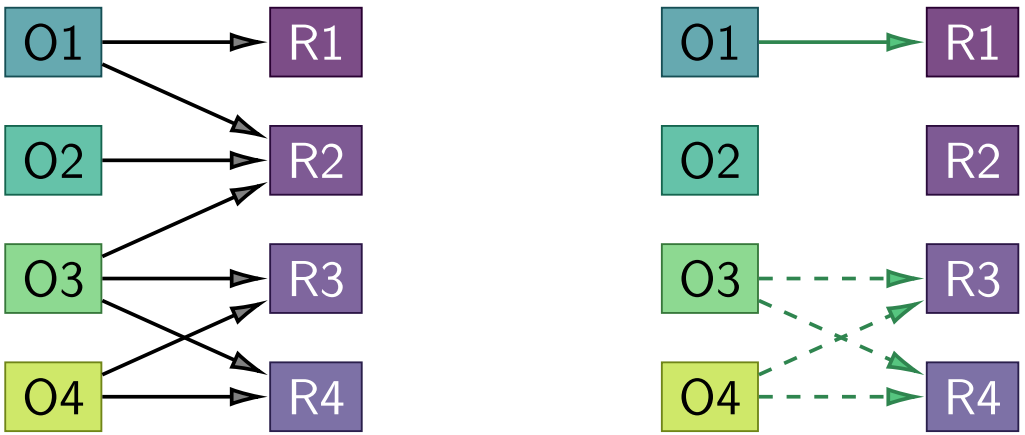
- Outputs O1-O4 are referenced in rings R1-R4

Zero Mixin Removal & Intersection Removal



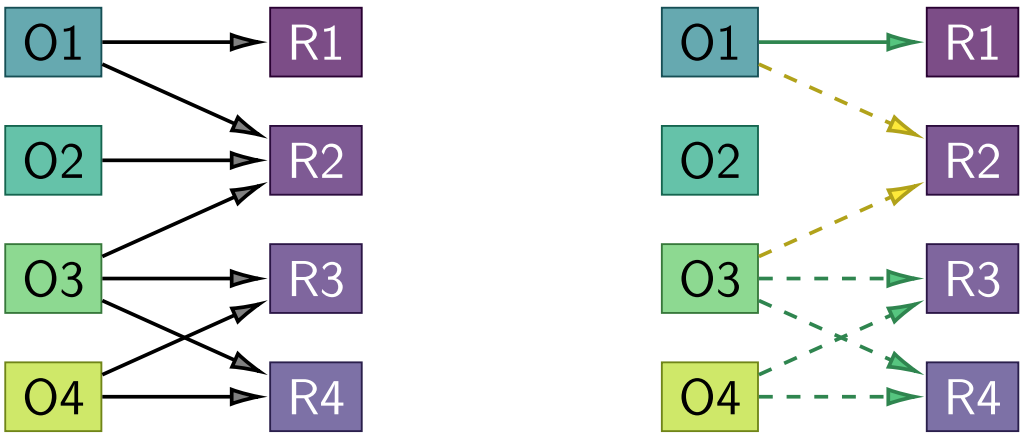
- Outputs O1-O4 are referenced in rings R1-R4
- R1 only references O1 $\implies$ must be the real input

Zero Mixin Removal & Intersection Removal



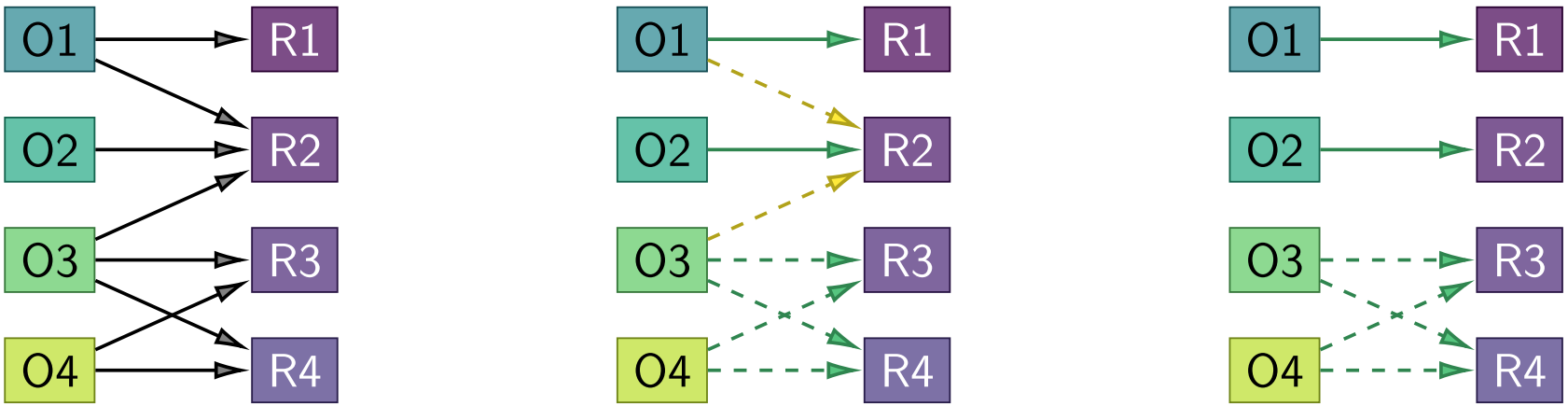
- Outputs O1-O4 are referenced in rings R1-R4
- R1 only references O1 $\implies$ must be the real input
- $I = \{R3, R4\}$ reference $O = \{O3, O4\}$
  $|I| = |O| \implies$ O3 & O4 spent in R3 & R4

Zero Mixin Removal & Intersection Removal



- Outputs O1-O4 are referenced in rings R1-R4
- R1 only references O1 $\implies$ must be the real input
- $I = \{R3, R4\}$ reference $O = \{O3, O4\}$
  $|I| = |O| \implies$ O3 & O4 spent in R3 & R4
- $R2$ only has one non-mixin reference remaining.

Zero Mixin Removal & Intersection Removal



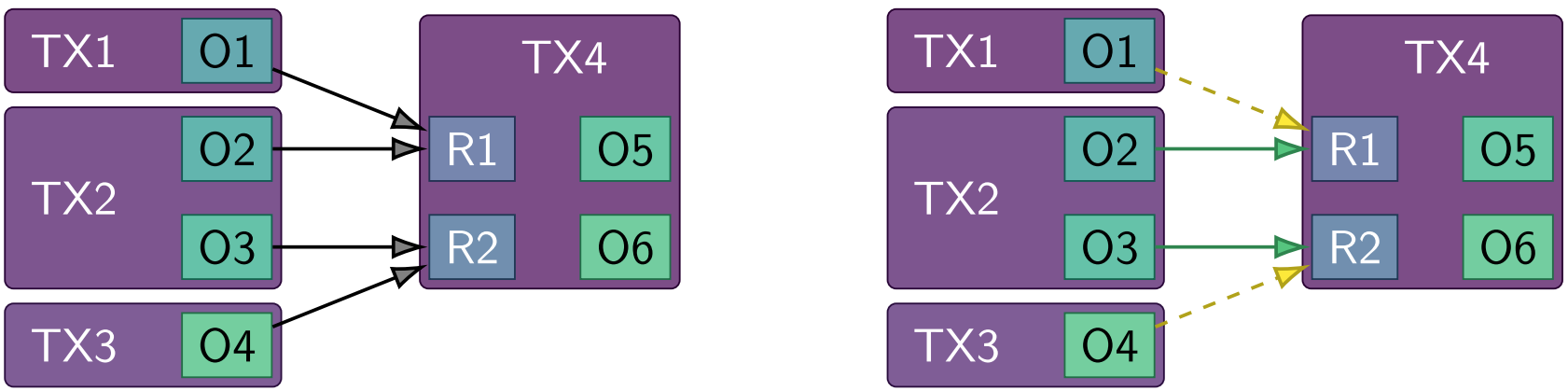
- Outputs O1-O4 are referenced in rings R1-R4
- R1 only references O1 $\implies$ must be the real input
- $I = \{R3, R4\}$ reference $O = \{O3, O4\}$
  $|I| = |O| \implies$ O3 & O4 spent in R3 & R4
- $R2$ only has one non-mixin reference remaining.

Output Merging Heuristic (OMH)

- Output merging mostly due to denomination splitting:
  - Initially, amounts were disclosed on blockchain
  - Ring signatures required multiple outputs with identical amounts
  - Outputs were partitioned to facilitate this ($7 \rightarrow 5 + 2$)

Output Merging Heuristic (OMH)

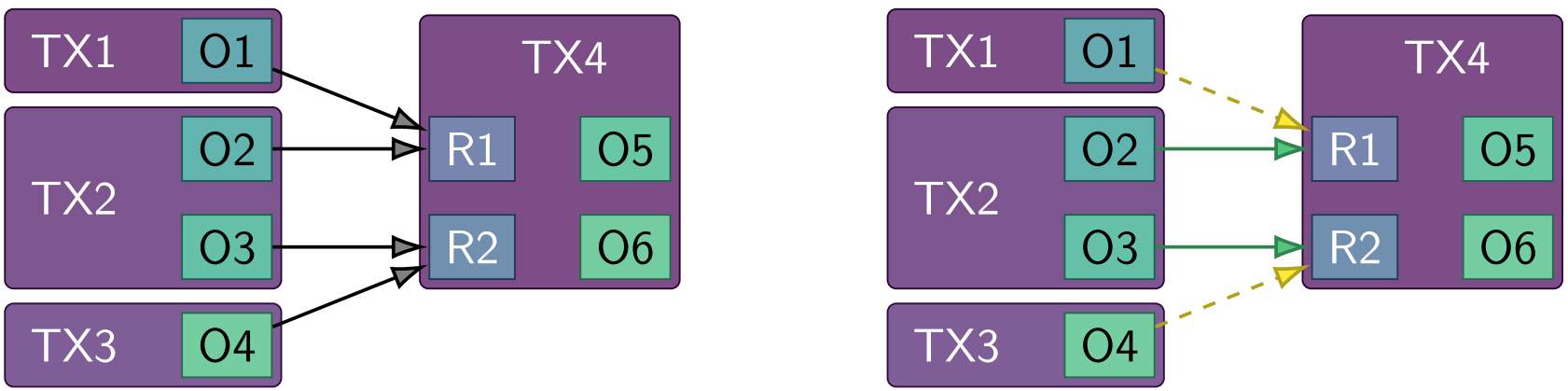
- Output merging mostly due to denomination splitting:
  - Initially, amounts were disclosed on blockchain
  - Ring signatures required multiple outputs with identical amounts
  - Outputs were partitioned to facilitate this ($7 \rightarrow 5 + 2$)



- TX4 has two inputs which reference a TXO from TX2

Output Merging Heuristic (OMH)

- Output merging mostly due to denomination splitting:
  - Initially, amounts were disclosed on blockchain
  - Ring signatures required multiple outputs with identical amounts
  - Outputs were partitioned to facilitate this ($7 \rightarrow 5 + 2$)



- TX4 has two inputs which reference a TXO from TX2
- OMH assumes that these outputs are real

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs
- TX outputs were partitioned into denominations ($7 \rightarrow 5 + 2$)

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs
- TX outputs were partitioned into denominations ($7 \rightarrow 5 + 2$)
  - $\rightarrow$ guessing real inputs mostly correct

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs
- TX outputs were partitioned into denominations ($7 \rightarrow 5 + 2$)
  - $\rightarrow$ guessing real inputs mostly correct
- Temporal distribution of decoys and spent outputs don't match

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs
- TX outputs were partitioned into denominations ($7 \rightarrow 5 + 2$)
  - $\rightarrow$ guessing real inputs mostly correct
- Temporal distribution of decoys and spent outputs don't match
  - $\rightarrow$ Educated guessing based on output-age effective

Traceability Analysis

[Kumar et al., 2017] and [Möser et al., 2018] found:

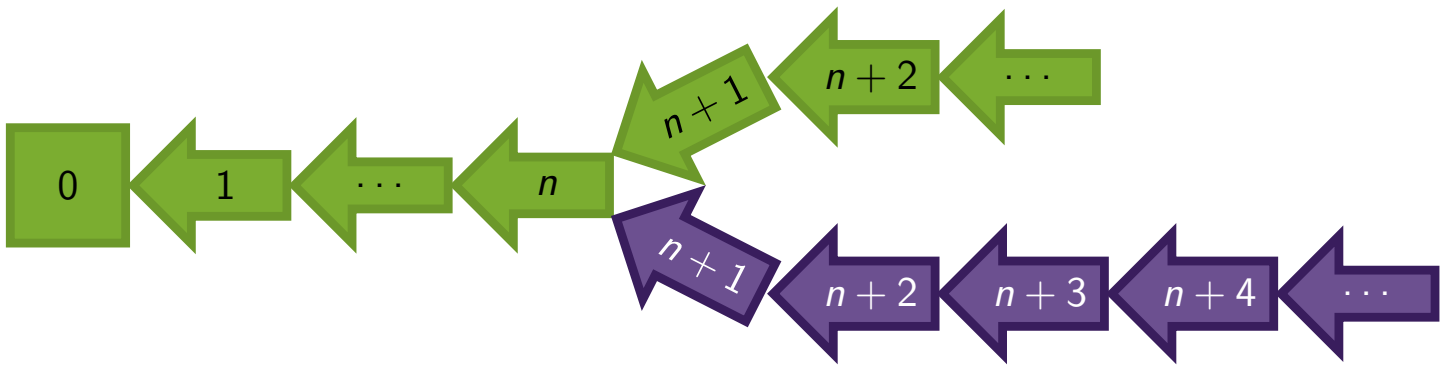
- Iteratively removing known spent outputs from rings allows identification of new spent outputs
  - $\rightarrow$ identified majority of real spent outputs
- TX outputs were partitioned into denominations ($7 \rightarrow 5 + 2$)
  - $\rightarrow$ guessing real inputs mostly correct
- Temporal distribution of decoys and spent outputs don't match
  - $\rightarrow$ Educated guessing based on output-age effective
- Transaction protocol has been updated since

Improvements to the protocol

- ZMR works like a chain reaction from an initial set of inputs without decoys.
  - Since 2016, the mandatory minimum ringsize has been increased
  - Minimum ringsizes + RingCT TX were effective
  - Ringsize $\equiv 11$ since last update
- Mixin sampling has been improved with different approaches
  - Triangular distribution
  - Recent zone: Force 25-50% recent outputs
  - Gamma distribution: Distribution based on empirical analysis

Currency hardforks

- A cryptocurrency can be forked, resulting in two currencies.
- A fork can either start a new blockchain or continue the existing chain.



- Pre-fork funds can be spent on both chains
- This can be exploited for linking/tracing analysis

Cross Chain (Fork) Analysis

- Double spends are prevented with *key images*

Cross Chain (Fork) Analysis

- Double spends are prevented with *key images*
- Key image is derived from spent output and may occur at most once on the TX record

Cross Chain (Fork) Analysis

- Double spends are prevented with *key images*
- Key image is derived from spent output and may occur at most once on the TX record
- Method to derive key image must be identical on all branches

Cross Chain (Fork) Analysis

- Double spends are prevented with *key images*
- Key image is derived from spent output and may occur at most once on the TX record
- Method to derive key image must be identical on all branches
- If two rings on two branches have the same key image, they spend the same TXO.

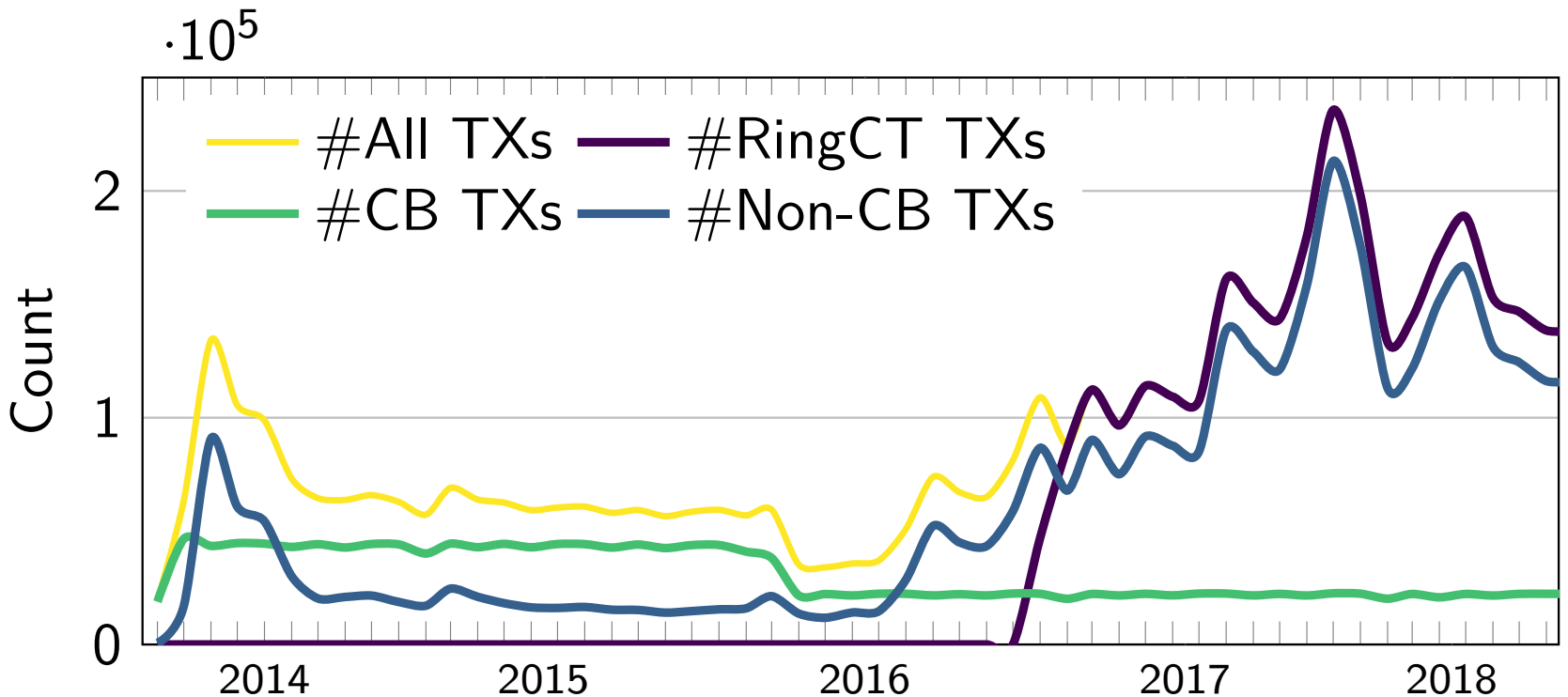
Contribution of this work

- Up to date evaluation of existing methods
  - Previous studies published shortly after introduction of RingCT
  - Changes to mixin sampling and ringsize in 09/2017 and 04/2018.
- Quantify impact on traceability from recent Monero hardforks
  - Monero Original: Continuation of Monero v6 (ASIC compatible)
  - MoneroV: Fork with some changes to emission curve

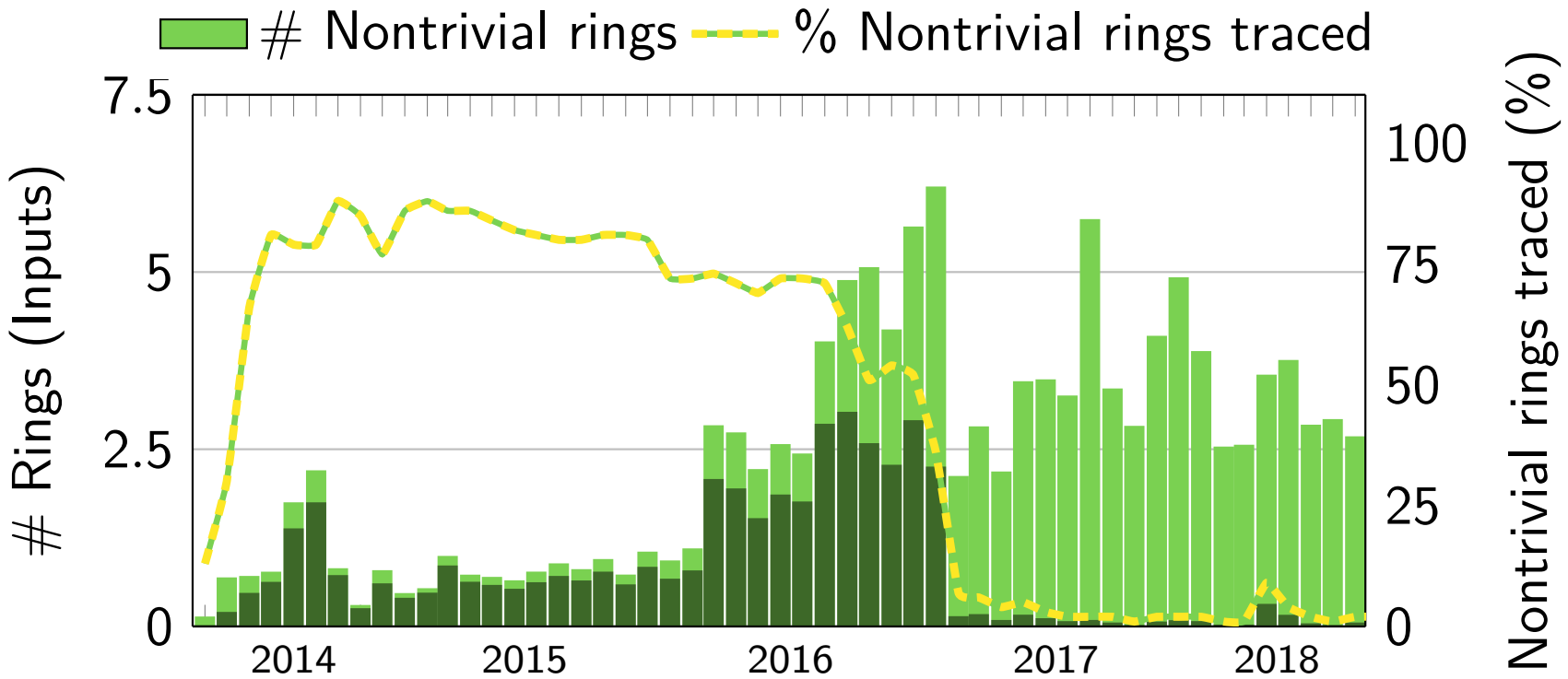
Dataset & Method

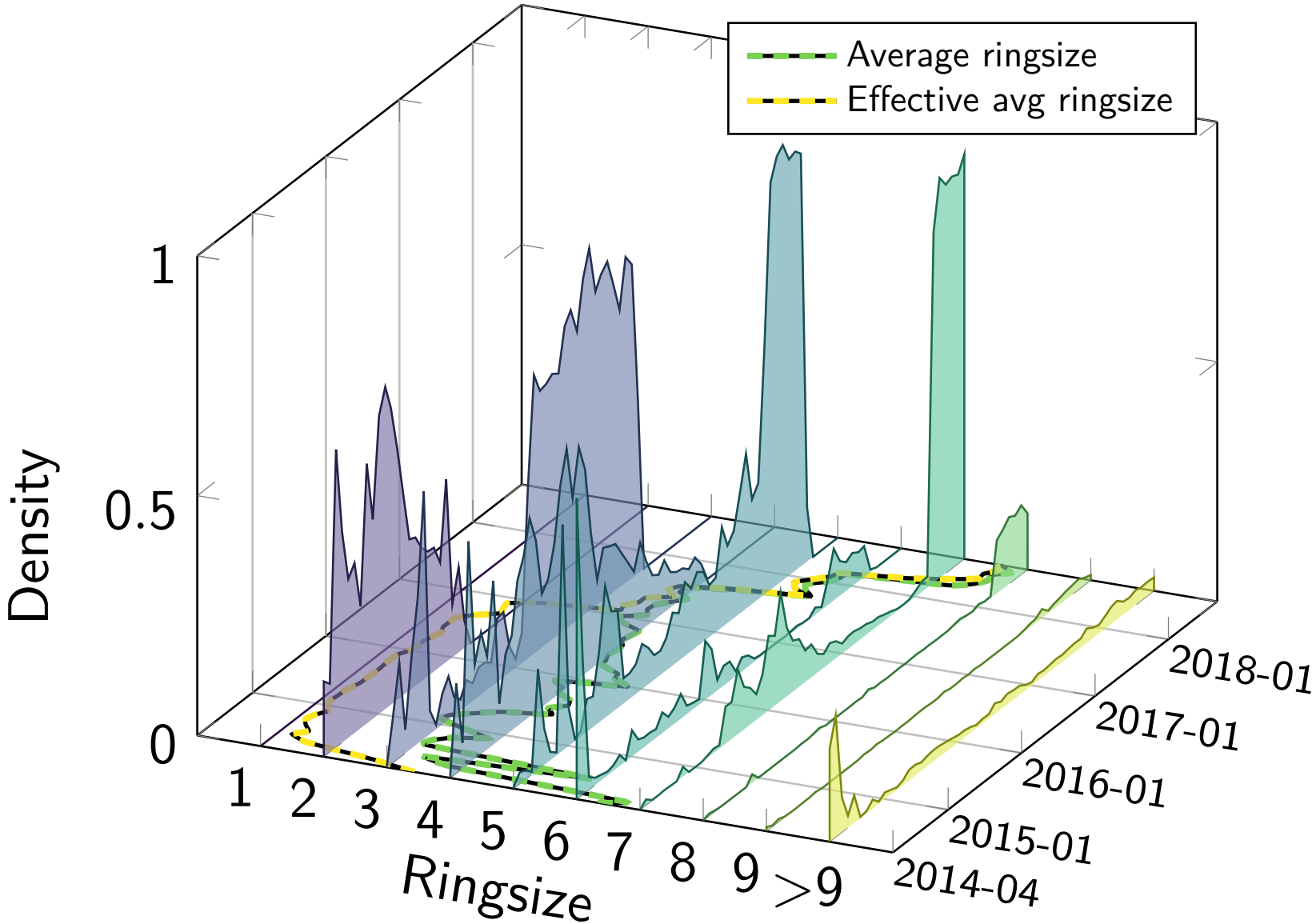
1. Exported Monero (XMR), MoneroV (XMV) and Monero Original (XMO) blockchain up to Aug. 31th, 2018.
2. Employed Zero Mixin Removal & Intersection Removal
3. Added fork data and applied cross chain analysis (+ZMR/IR)
4. Applied heuristics from [Kumar et al., 2017] and [Möser et al., 2018]:
   - Guess Newest Heuristic
   - Output Merging Heuristic
5. Evaluated accuracy with ground truth (where possible) with results from steps 3.

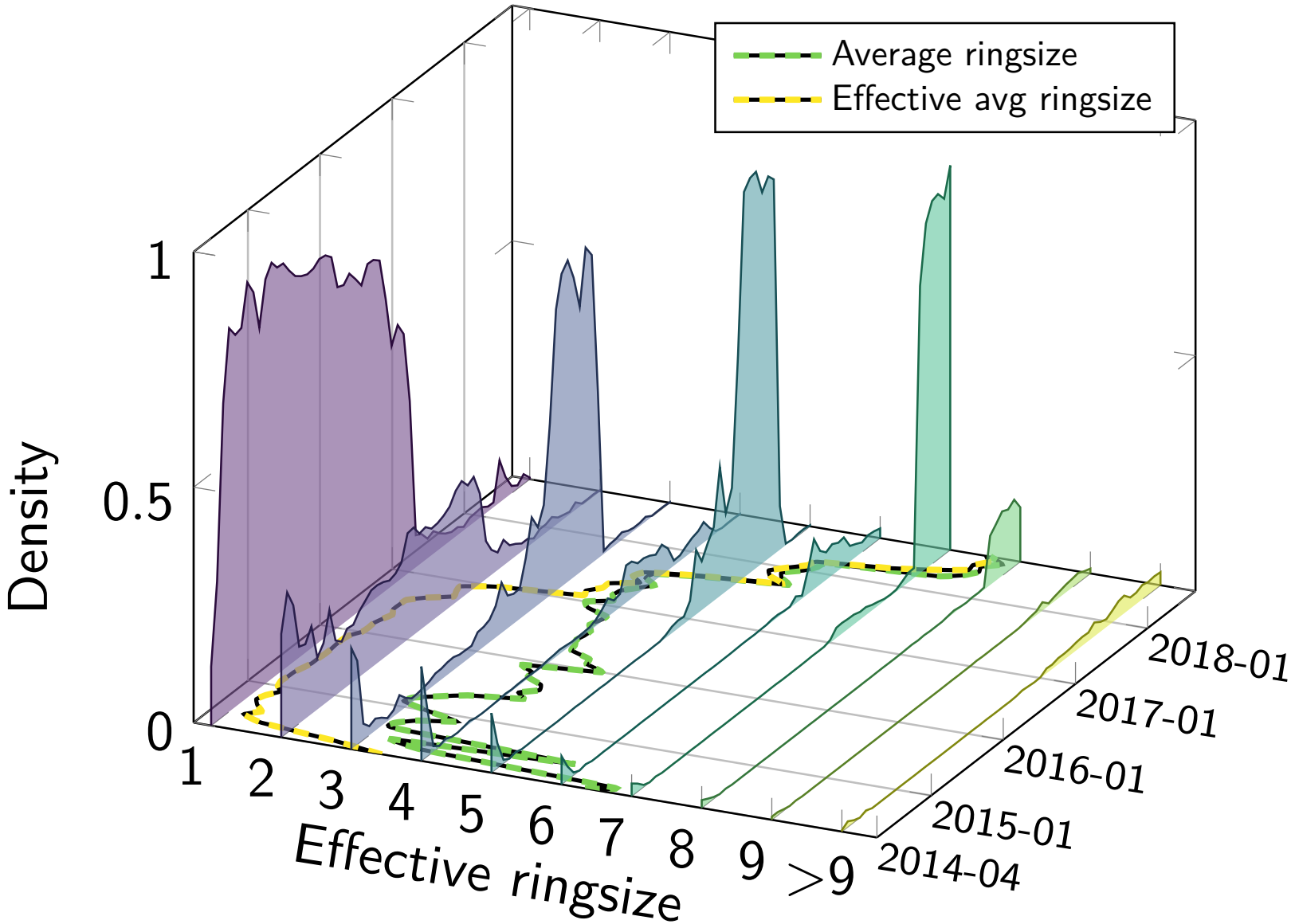
Monero Activity



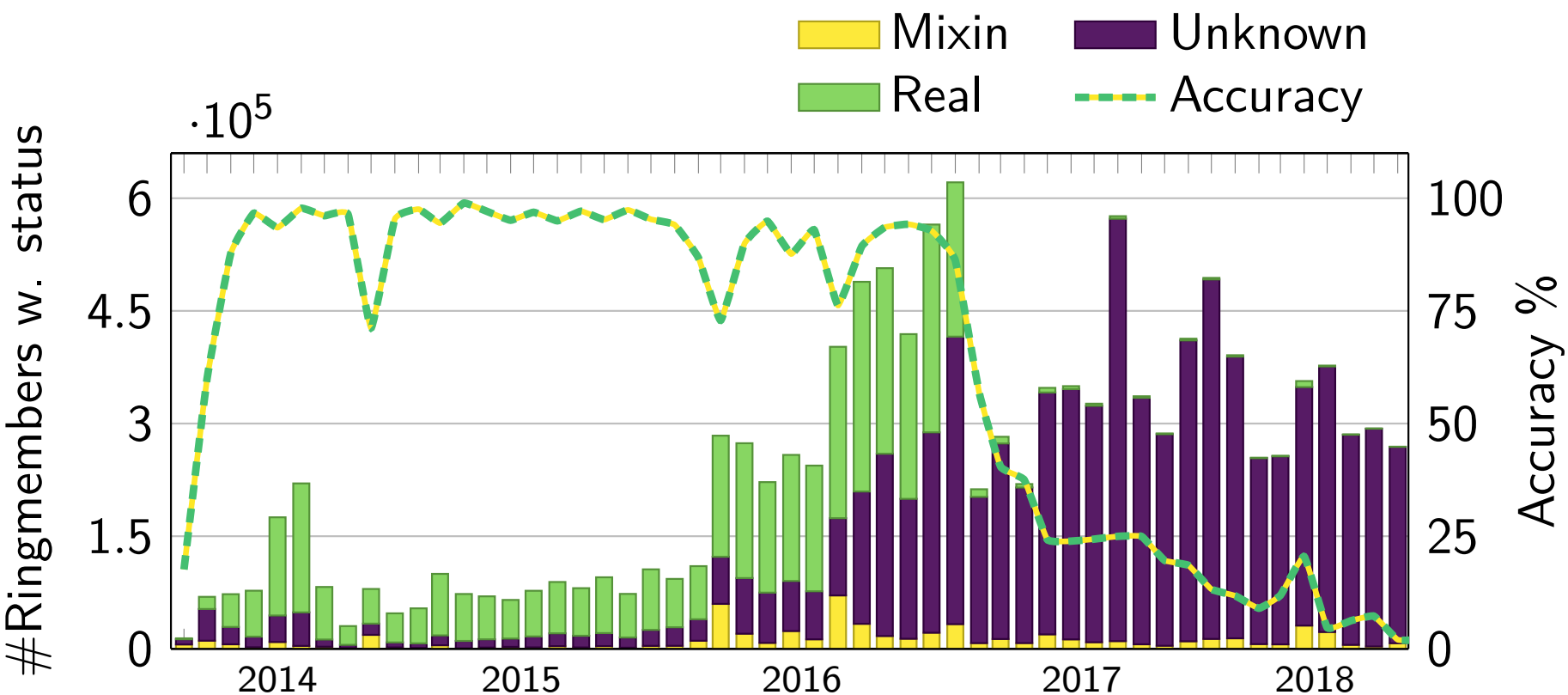
Traced Inputs



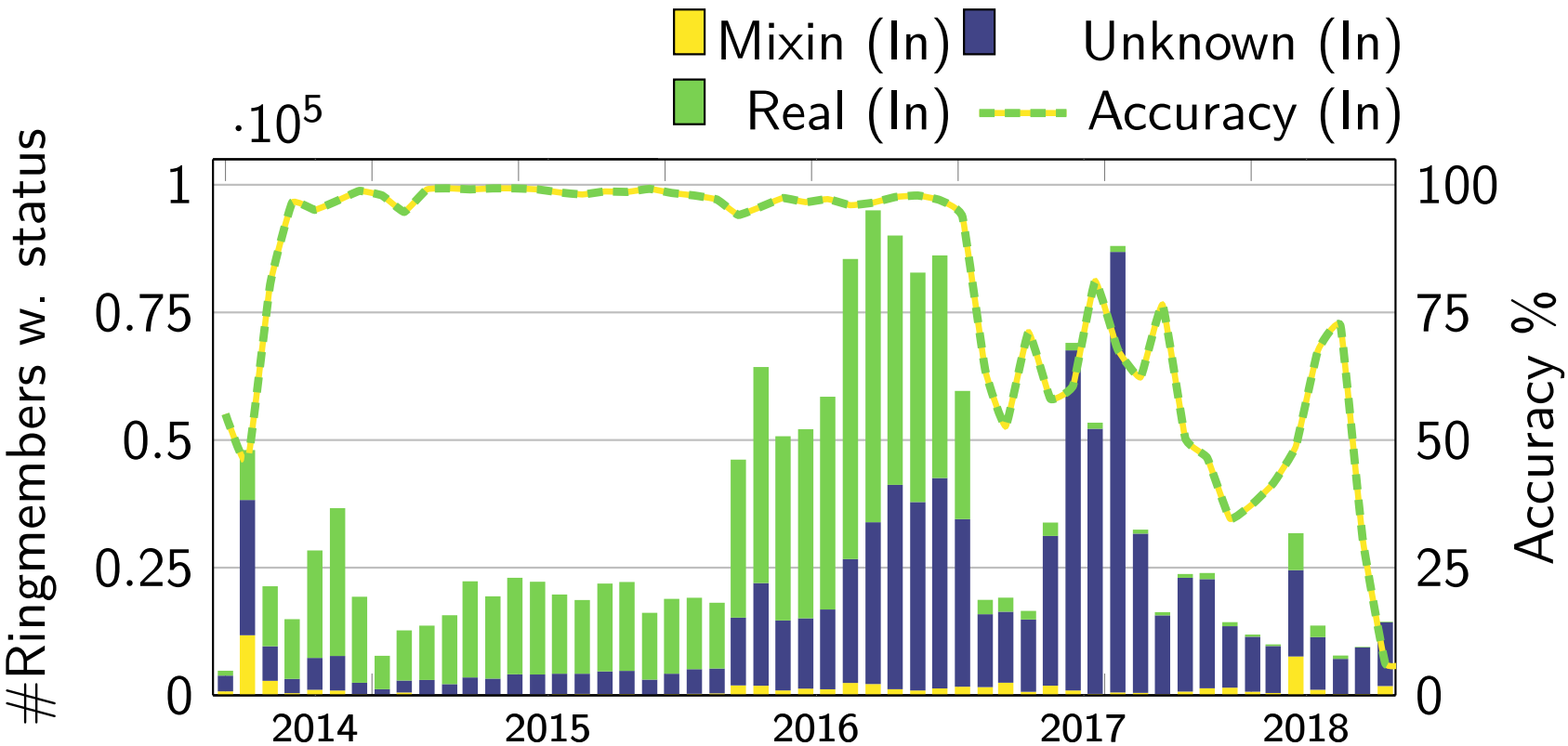

Average ringsize
Effective avg ringsize
Density
1
0.5
0
1 2 3 4 5 6 7 8 9 >9
Ringsize
2014-04
2015-01
2016-01
2017-01
2018-01


Average ringsize
Effective avg ringsize
Density
1
0.5
0
1
2
3
4
5
6
7
8
9
>9
Effective ringsize
2014-04
2015-01
2016-01
2017-01
2018-01

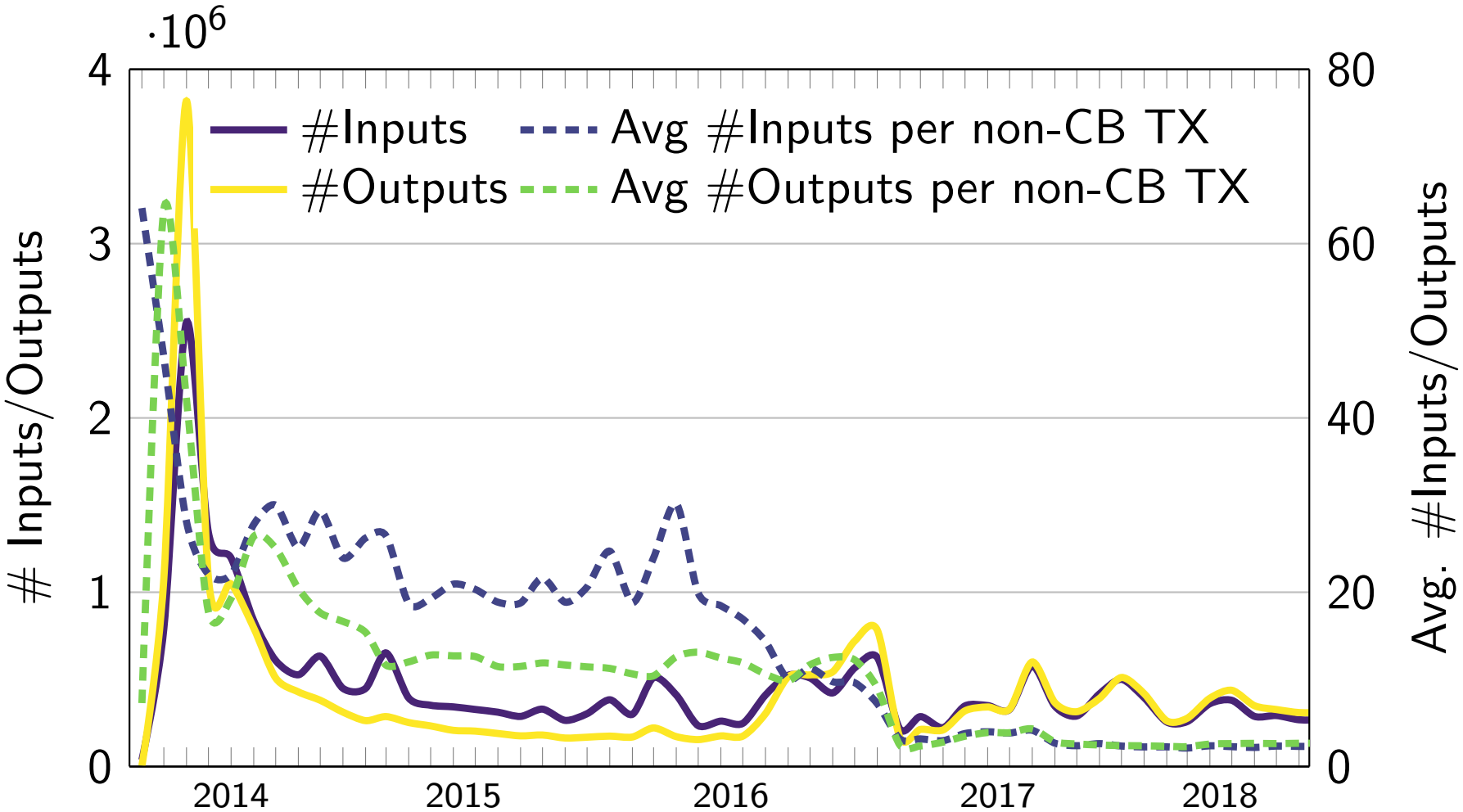
Guess Newest Heuristic



Output Merging Heuristic



Inputs/Outputs (per TX)



Summary

- Nowadays, most Monero TXs are untraceable

Summary

- Nowadays, most Monero TXs are untraceable
- Guess Newest Heuristic does not work with current mixin sampling technique

Summary

- Nowadays, most Monero TXs are untraceable
- Guess Newest Heuristic does not work with current mixin sampling technique
- Impact from Cross Chain Analysis not very large

Summary

- Nowadays, most Monero TXs are untraceable
- Guess Newest Heuristic does not work with current mixin sampling technique
- Impact from Cross Chain Analysis not very large
  1. Forks didn't have a lot of traction

Summary

- Nowadays, most Monero TXs are untraceable
- Guess Newest Heuristic does not work with current mixin sampling technique
- Impact from Cross Chain Analysis not very large
  1. Forks didn't have a lot of traction
  2. Mandatory ringsize of 7 enough to prevent chain reactions

References I


Kumar, A. et al. (2017).
A traceability analysis of Monero's blockchain.
In *European Symposium on Research in Comp. Sec.*, pages 153–173. Springer.


Möser, M. et al. (2018).
An Empirical Analysis of Traceability in the Monero Blockchain.
PoPET, 2018(3):143–163, DOI: 10.1515/popets-2018-0025, https://content.sciendo.com/view/journals/popets/2018/3/article-p143.xml.


Nakamoto, S. (2008).
Bitcoin: A peer-to-peer electronic cash system.

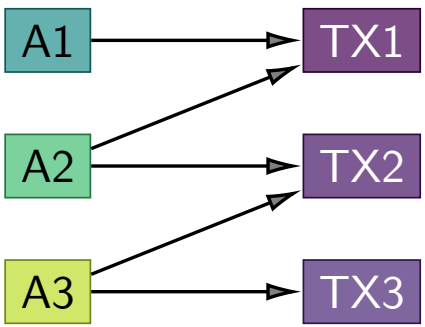
References II


Van Saberhagen, N. (2013).
Cryptonote v 2. 0.
https://cryptonote.org/whitepaper.pdf,
https://cryptonote.org/whitepaper.pdf.

Bitcoin analytics

- Analysis techniques impeded privacy of Bitcoin
- Sets of addresses belonging to a user can often be identified
  - Multi Input Heuristic
  - Change Heuristics
- Simplified transaction graph allows further analysis

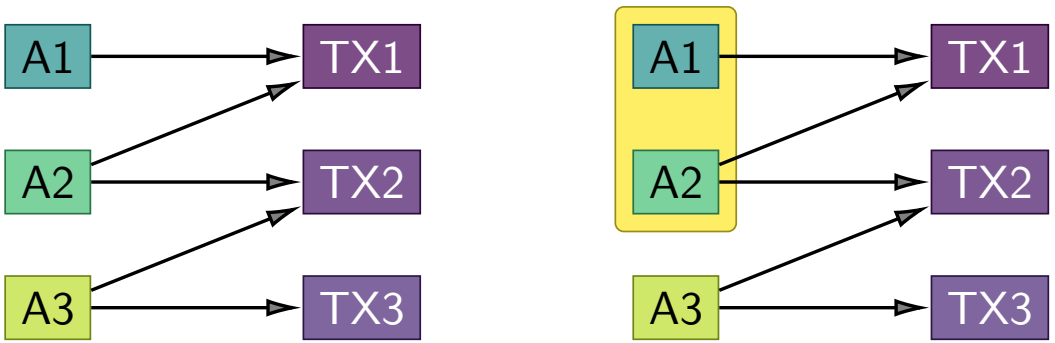
Multiple Input Heuristic (MIH)



- Transactions TX1-TX3 spend outputs belonging to A1-A3

[0]Back: 5

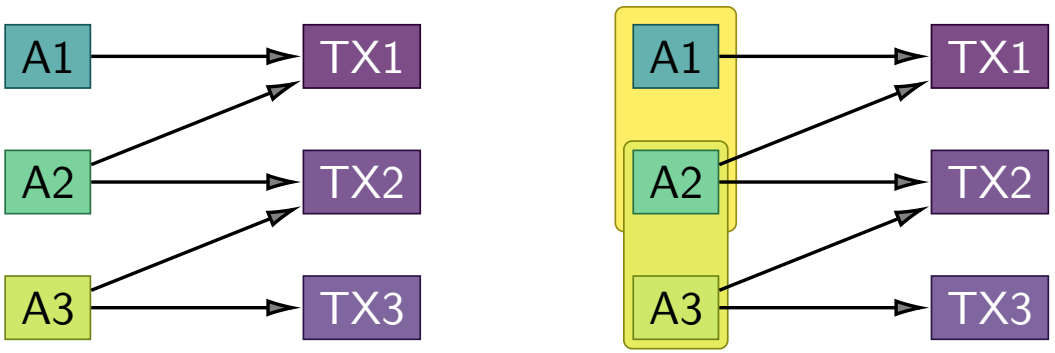
Multiple Input Heuristic (MIH)



- Transactions TX1-TX3 spend outputs belonging to A1-A3
- All outputs spent in a TX likely belong to the same person

[0]Back: 5

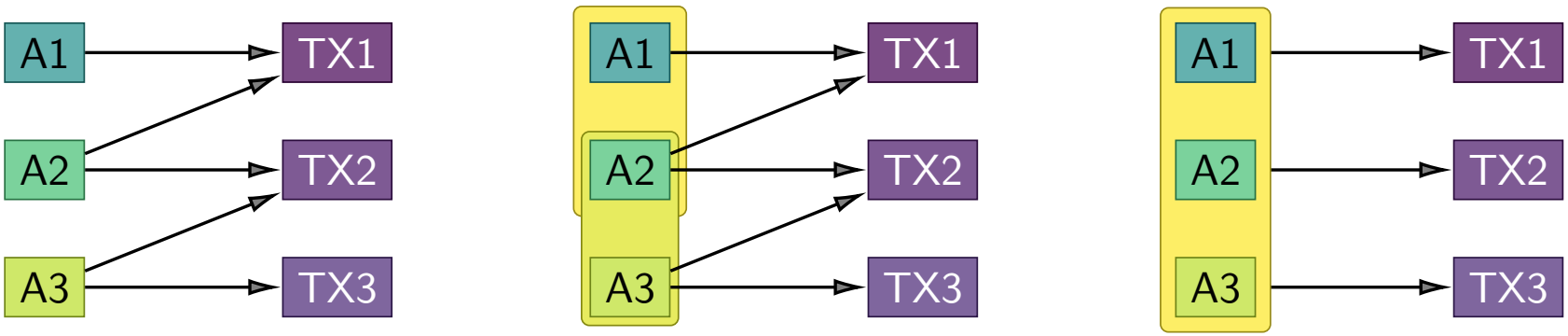
Multiple Input Heuristic (MIH)



- Transactions TX1-TX3 spend outputs belonging to A1-A3
- All outputs spent in a TX likely belong to the same person
- Overlapping clusters are merged

[0]Back: 5

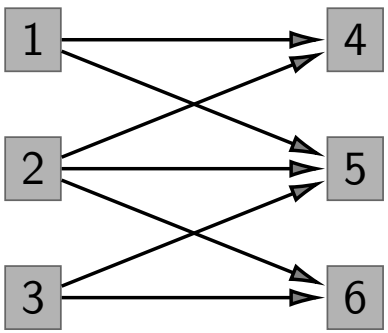
Multiple Input Heuristic (MIH)



- Transactions TX1-TX3 spend outputs belonging to A1-A3
- All outputs spent in a TX likely belong to the same person
- Overlapping clusters are merged

[0]Back: 5

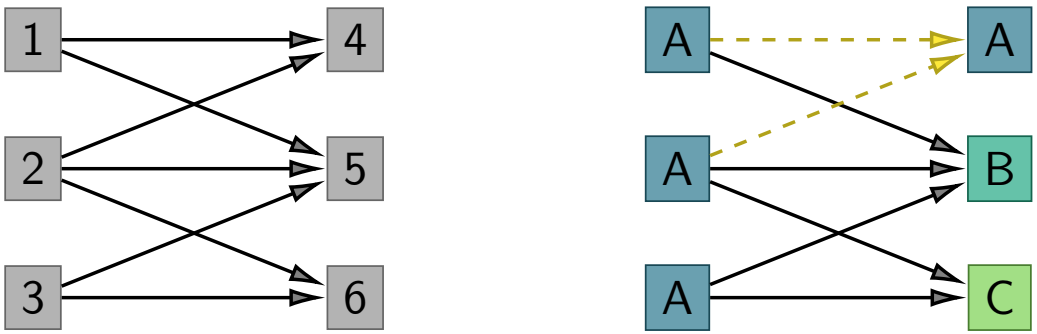
Address Clustering



- Address graph shows addresses (1-6) and TXs (edges)

[0]Back: 5

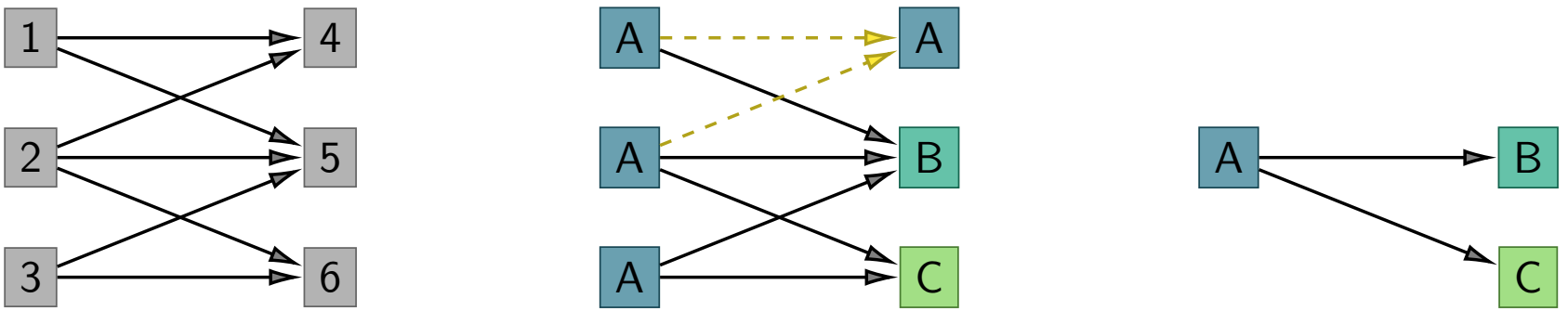
Address Clustering



- Address graph shows addresses (1-6) and TXs (edges)
- Use information from e.g. MIH to label nodes

[0]Back: 5

Address Clustering



- Address graph shows addresses (1-6) and TXs (edges)
- Use information from e.g. MIH to label nodes
- Simplify graph: Address graph $\Rightarrow$ Entity graph

[0]Back: 5