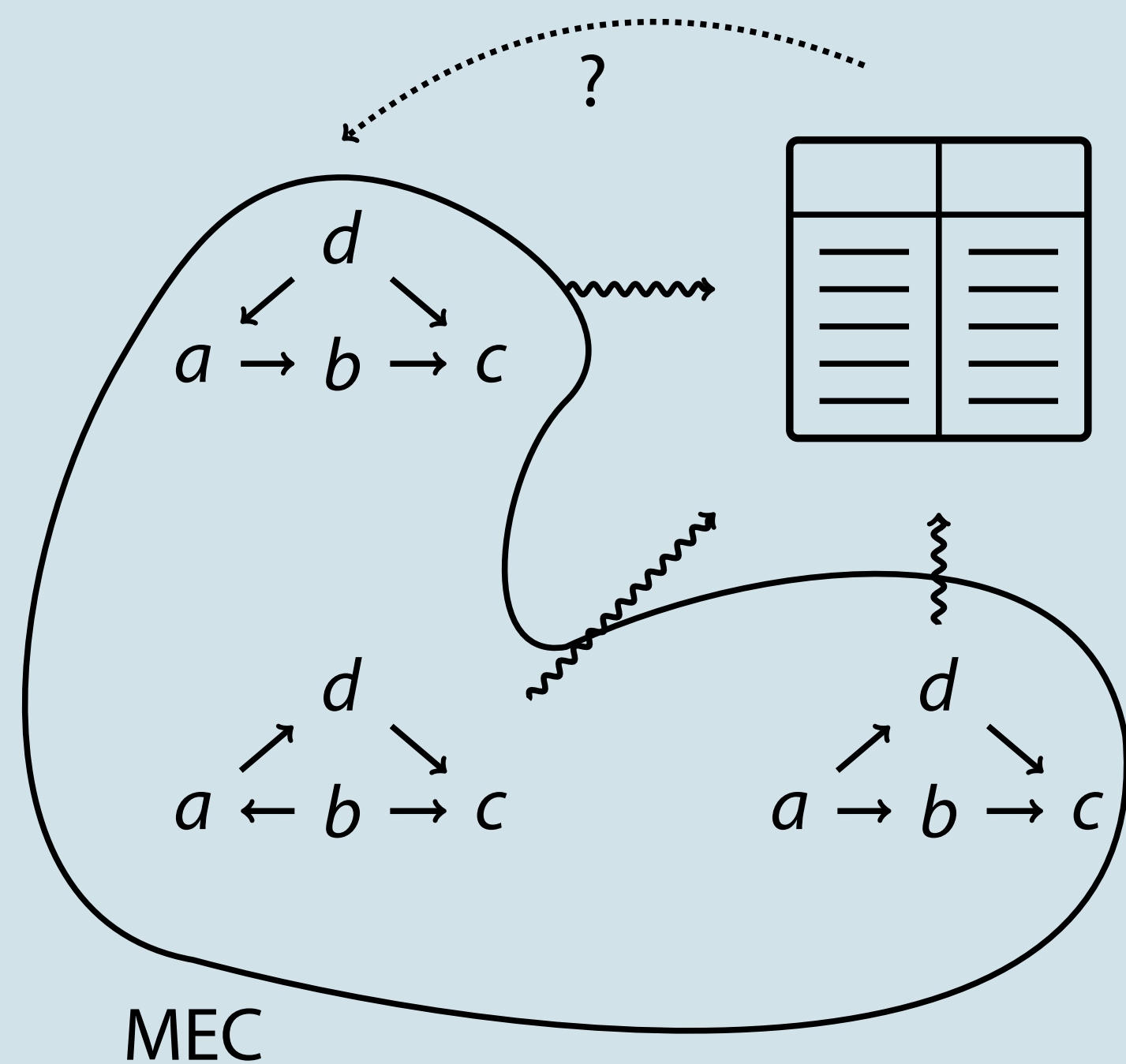




Efficient Enumeration of Markov Equivalent DAGs

Marcel Wienöbst, Malte Luttermann, Max Bannach and Maciej Liśkiewicz

1. Abstract



Fundamental problem of causal discovery:
Multiple causal structures may generate the same observations.

Those with the same statistical properties form a Markov equivalence class (MEC).

Problem: Enumerate all directed acyclic graphs (DAGs) in an MEC efficiently.

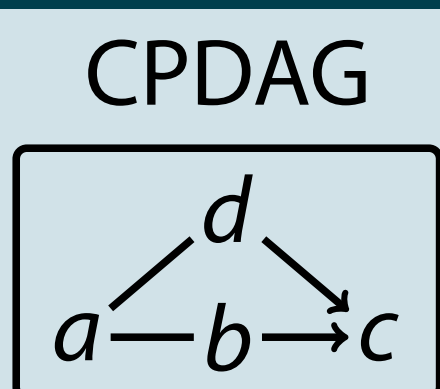
Main Result: Linear-time delay algorithm for enumerating the DAGs in an MEC.

2. Enumeration of Markov equivalence classes (MECs)

Input: Compact representation of an MEC as CPDAG.

Task: List all DAGs in MEC one-by-one.

Objective: Small *delay* between successive outputs.



	Approach	Delay
Meek '95	Meek-Rule Recursion	$O(m \cdot \text{meek}(n, m))$
Chickering '95	Transformational MEC Char.	$O(m^3)$
This work	Max. Cardinality Search (MCS)	$O(n + m)$

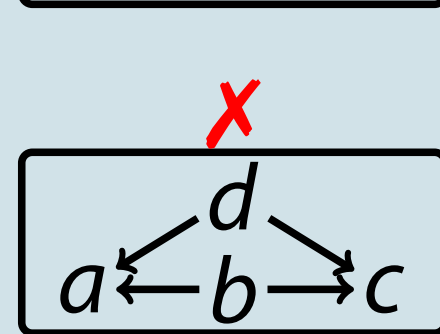
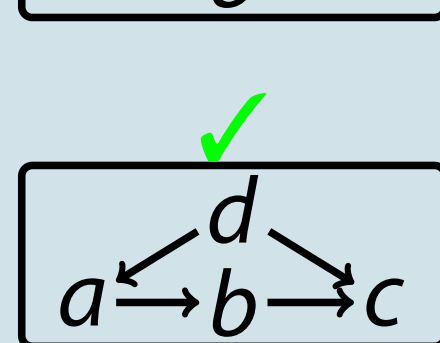
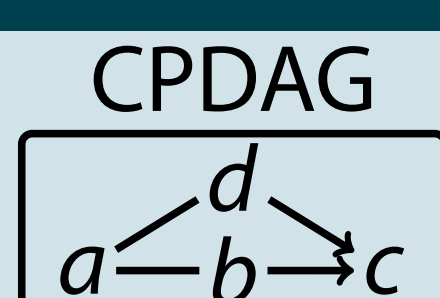
3. Prerequisite: Understanding Extension of CPDAGs

Input: Compact representation of an MEC as CPDAG.

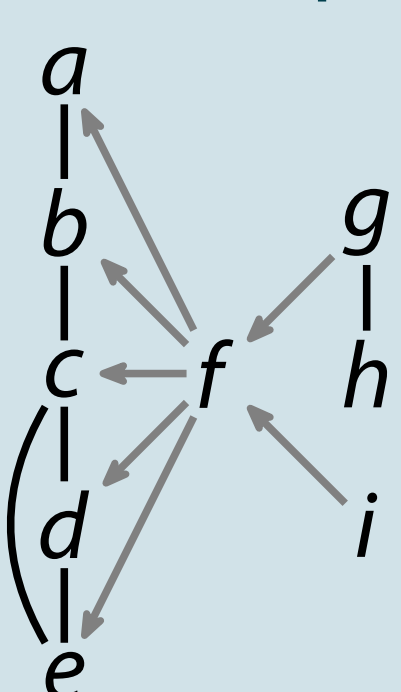
Problem: Compute any DAG in the MEC.

Algorithm: Folklore $O(n + m)$ approach.

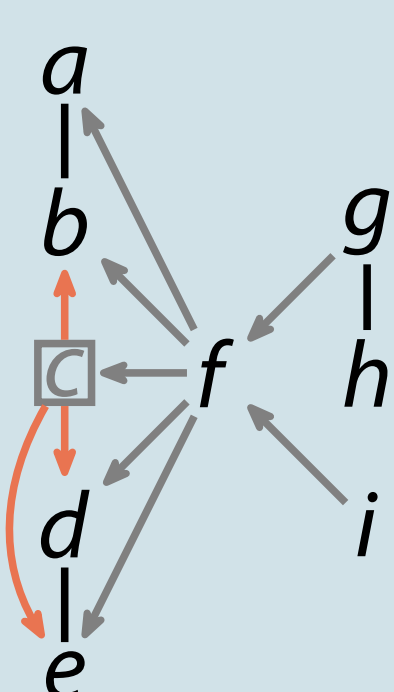
- Discard all oriented edges.
- Find acyclic orientation without v-structure using MCS (traverse vertices by highest number of visited neighbors).



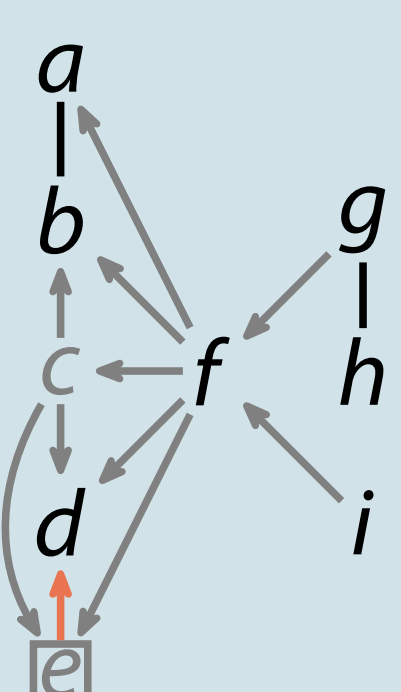
After Step (i)



Step (ii).1



Step (ii).2



vis. neighbors:
a b c d e f g h i
0 0 0 0 0 0 0 0 0

vis. neighbors:
a b c d e f g h i
0 1 0 1 1 0 0 0 0

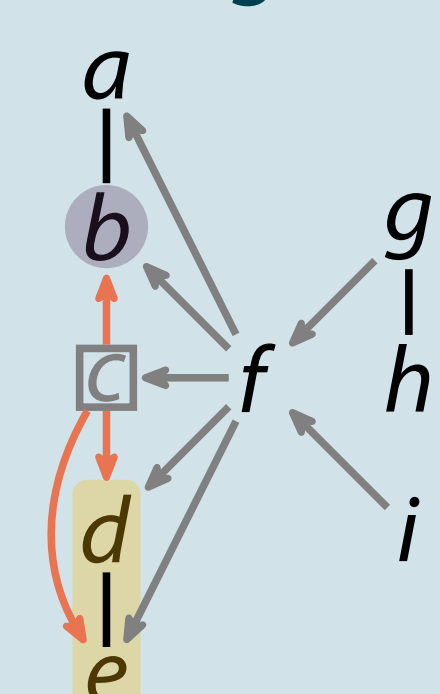
vis. neighbors:
a b c d e f g h i
0 1 0 2 1 0 0 0 0

4. Enumeration of an MEC with Linear-Time Delay

Idea: Recursively try all vertices with highest number of visited neighbors.

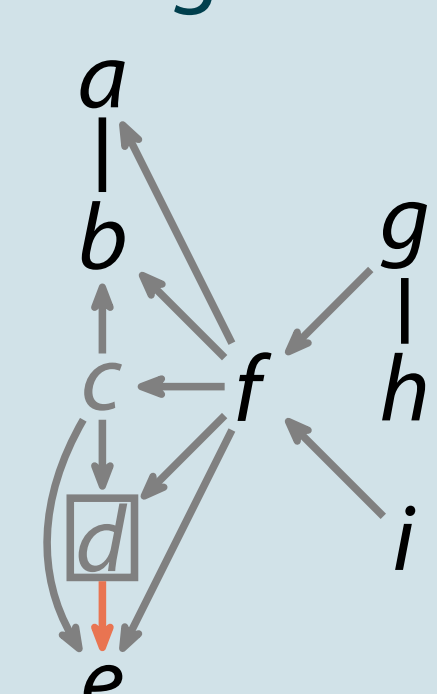
Problem: Possibly duplicate outputs.

Visiting c first:



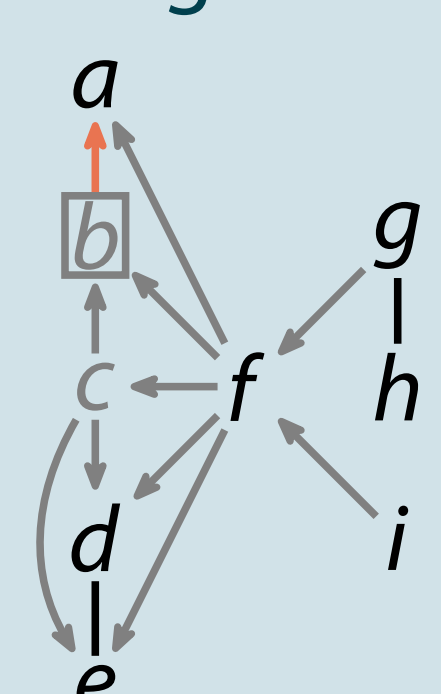
vis. neighbors:
a b c d e f g h i
0 1 0 1 1 0 0 0 0

Visiting d after c:



vis. neighbors:
a b c d e f g h i
0 1 0 1 2 0 0 0 0

Visiting b after c:



vis. neighbors:
a b c d e f g h i
1 1 0 1 1 0 0 0 0

Lemma: (Avoiding Duplicate Outputs)

If vertices x and y are...

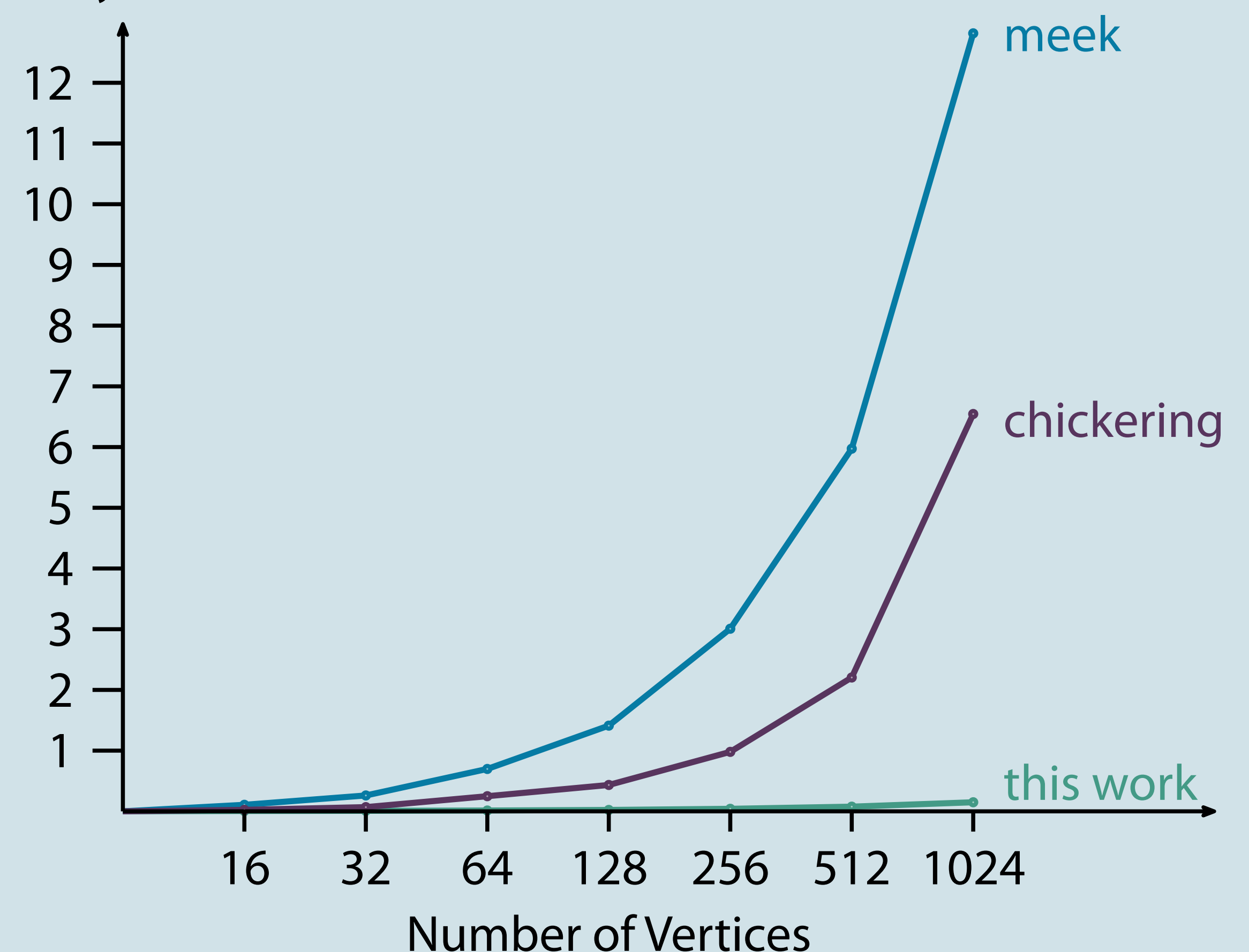
- ...connected: choosing either results in disjoint extensions.
- ...unconnected: any extension produced by choosing x first, can also be produced by choosing y first.

Theorem: (Main Result)

- An MEC can be enumerated with delay $O(n + m)$.
- For background knowledge, an $O(n^3)$ initialization step is needed, subsequent delay is $O(n + m)$.

5. Experimental Evaluation (sparse undirected graphs)

Delay in ms



6. Structural Insights Into Markov Equivalence

Theorem: (On the structure of MEC enumeration)

Every Markov equivalence class can be enumerated such that successive DAGs have structural hamming distance (that is, number of different edge orientations) less or equal than 3.