Rate-Agnostic (Causal) Structure Learning

S.M. Plis, D. Danks, C. Freeman, V. D. Calhoun

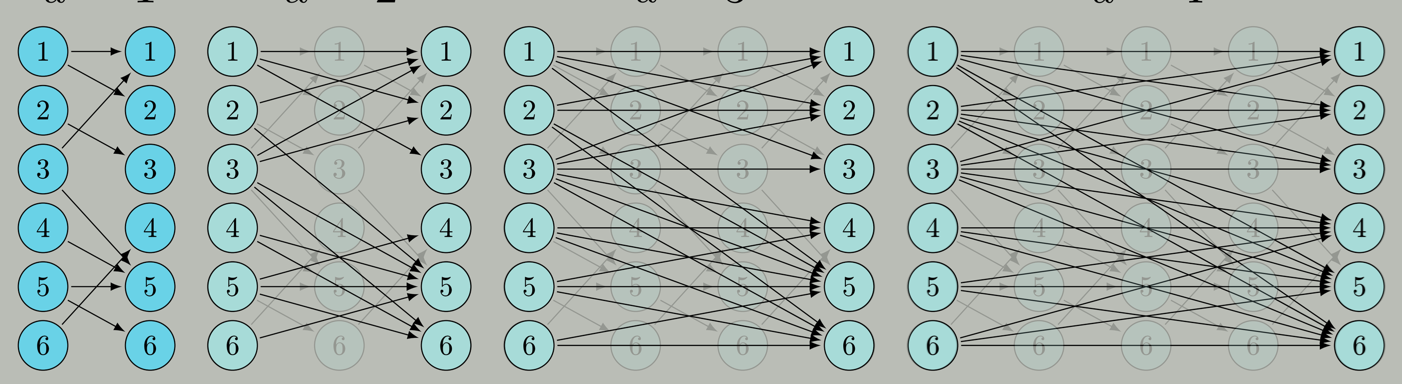
Carnegie Mellon University, The Mind Research Network

Abstract

- Causal structure learning from time series data is a major scientific challenge.
- Extant algorithms assume approximately equal system and measurement timescales.
- Often measurements occur at a significantly slower rate, but the size of the timescale mismatch is often unknown.
- We developed RASL (rate-agnostic structure learning) to learn structure from undersampled data.
- In simulated data, RASL algorithms reveal causal timescale structure and improved measurement timescale learning.

Undersampling Problem

- Measurement timescale significantly slower than causal or communication timescale



  - Intermediate time points are unobserved
  - The measurement timescale graph at undersampling rate u ($\mathcal{G}^u$) is different from the causal timescale graph $\mathcal{G}^1$
  - Denote with $\mathcal{H}$ a "compressed" representation of $\mathcal{G}^u$
- Given $\mathcal{G}^u$ for unknown u what can be inferred about $\mathcal{G}^1$?
  - Let $[\![\mathcal{H}]\!] = \{\mathcal{G}^1 : \exists u\, \mathcal{G}^u = \mathcal{H}\}$ be the equivalence class of $\mathcal{G}^1$ that could, for some undersample rate, yield $\mathcal{H}$.
  - Given $\mathcal{H}$ find $[\![\mathcal{H}]\!]$.
- Instrumental results:

Theorem

If $\mathcal{G}^u = \mathcal{G}^v$ for $u > v$, then $\forall w > u \exists k_w < u[\mathcal{G}^w = \mathcal{G}^{k_w}]$.

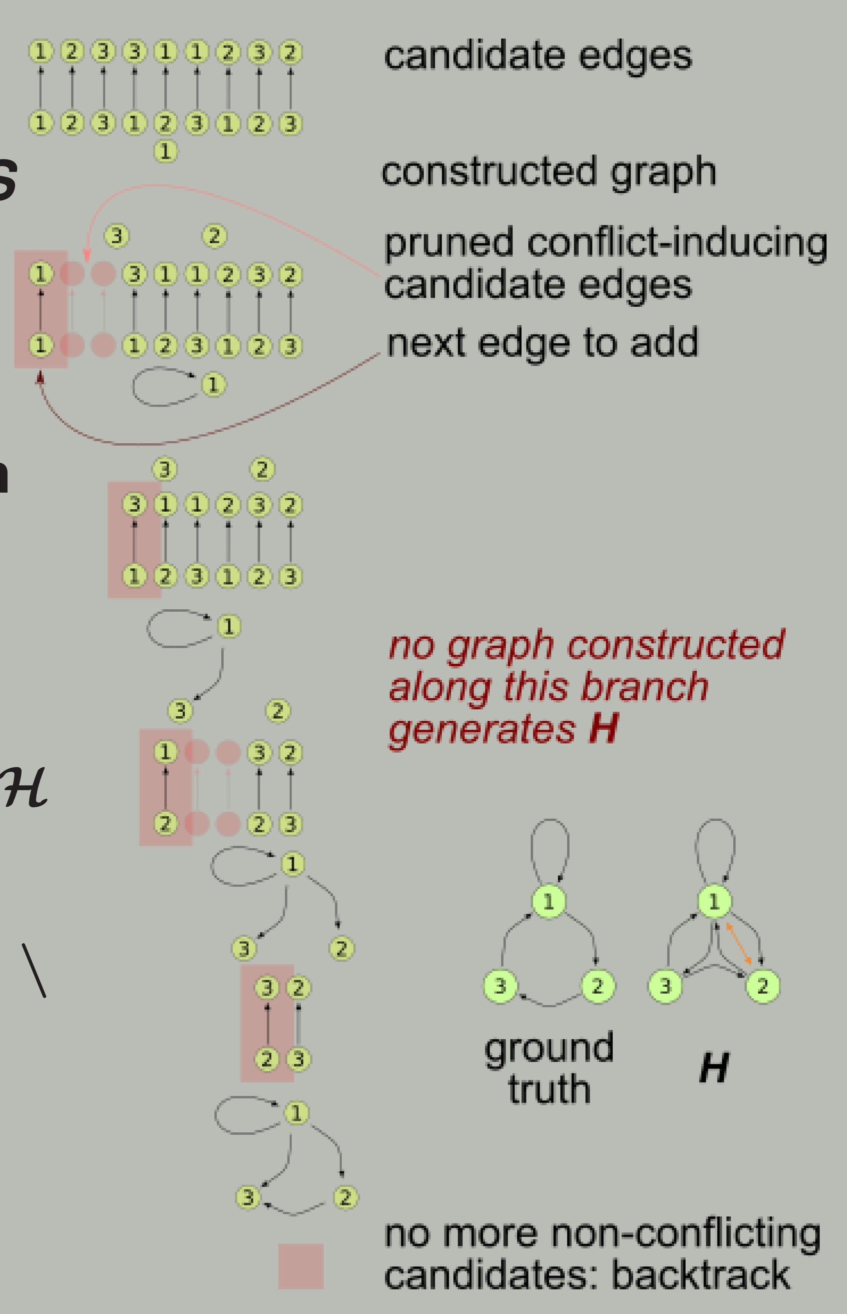
  - As u increases, if we find a graph that we previously encountered, then there cannot be any new graphs as $u \to \infty$
  - A (provable) stopping rule for some candidate $\mathcal{G}^1$: if $\mathcal{G}^u$ is not an edge-subset of $\mathcal{H}$ for all u, then do not consider any edge-superset of $\mathcal{G}^1$.

Recursive Algorithm

Algorithm *RecursiveEqClass*
Input: $\mathcal{H}$
Output: $[\![\mathcal{H}]\!]$
initialize empty graph $\mathcal{G}$ and set S

begin ***EdgeAdder*** $\mathcal{G}^*, \mathcal{H}, L$
  if *L has elements* **then**
    forall *the edges in L* **do**
      if *edge creates a conflict* **then**
        remove it from L
    if *L has elements* **then**
      forall *the edges in L* **do**
        add the edge to $\mathcal{G}^*$
        if $\exists \mathcal{G} \in \{(\mathcal{G}^*)^u\}$ *s.t.* $\mathcal{G} = \mathcal{H}$ **then**
          add $\mathcal{G}^*$ to S
        EdgeAdder $\mathcal{G}^*, \mathcal{H}, L \setminus$ the edge
        remove the edge from $\mathcal{G}^*$

put all n^2 edges into list L
EdgeAdder$(\mathcal{G}, \mathcal{H}, L)$
return S



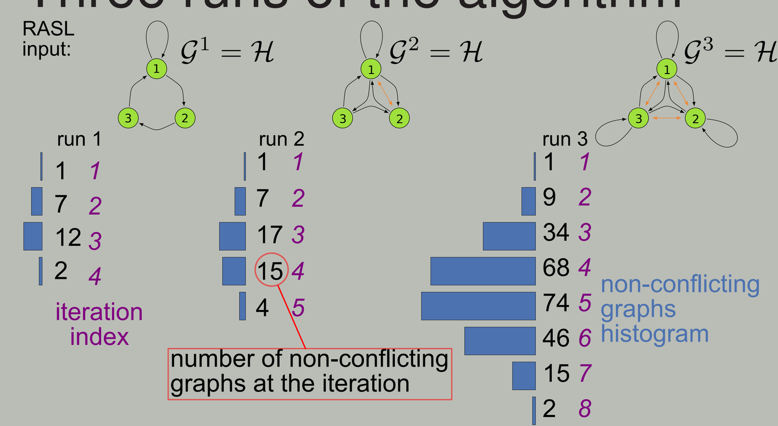
Start with an empty graph and n^2 possible edges. For each edge e, construct a graph $\mathcal{G}$ containing only e. If $\mathcal{G}^u \not\subseteq \mathcal{H}$ for all u, then reject; else if $\mathcal{G}^u = \mathcal{H}$ for some u, then add $\mathcal{G}$ to $[\![\mathcal{H}]\!]$; else, recurse into non-conflicting graphs in order.

Iterative Algorithm

Algorithm *IterativeEqClass*
Input: $\mathcal{H}$
Output: $[\![\mathcal{H}]\!]$

initialize empty sets S
init empty d and n^2 edges
while d **do**
  $d, S_i = NextIterGs(d, \mathcal{H})$
  $S = S \cup S_i$
return S

Three runs of the algorithm



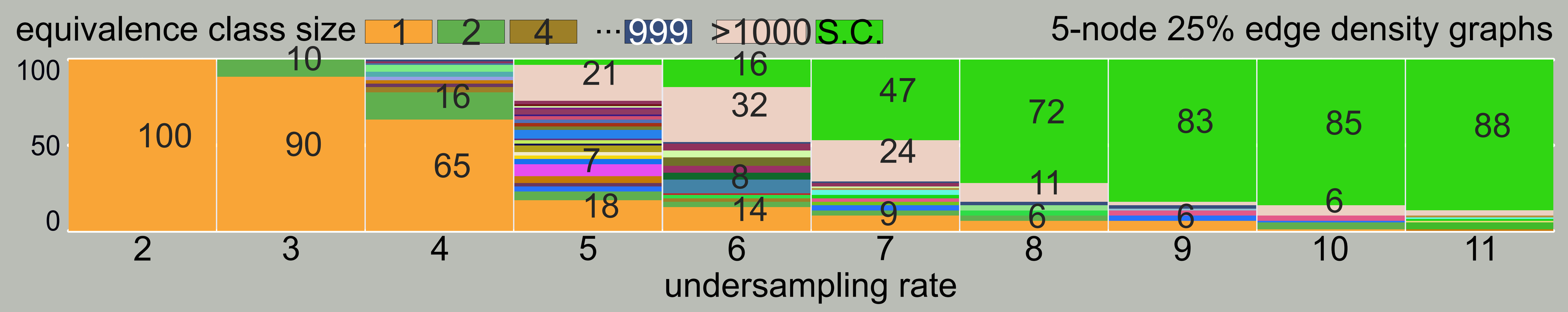
Procedure *NextIterGs*
Input: d structure, and $\mathcal{H}$
Output: d_r and set $S \subseteq [\![\mathcal{H}]\!]$

init empty d_r and S, S_i sets
forall *the graphs $\mathcal{G}$ in d* **do**
  forall *the edges e in $d(\mathcal{G})$* **do**
    if $e \notin \mathcal{G}$ **then**
      if *e conflicts with $\mathcal{G}$* **then**
        continue
      add e to $\mathcal{G}$
      if $\mathcal{G} \notin S_i$ **then**
        add $\mathcal{G}$ to S_i
        if *$\mathcal{G}$ conflicts with $\mathcal{H}$* **then**
          continue
        if $\exists \tilde{\mathcal{G}} \in \{\mathcal{G}^u\}$ *s.t.* $\tilde{\mathcal{G}} = \mathcal{H}$ **then**
          add $\mathcal{G}$ to S
      remove e from $\mathcal{G}$
  non-conflicting $\mathcal{G}$s w/ edges to d_r
return d_r, S

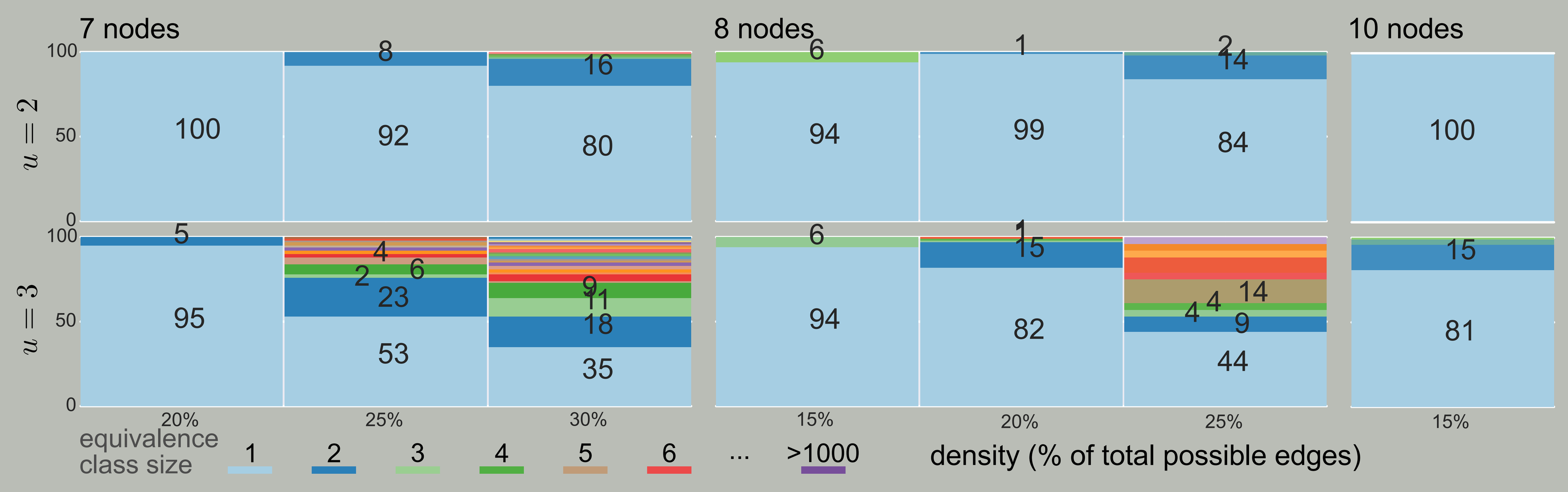
At stage $i+1$, RASL$_{ie}$ (a) considers each graph $\mathcal{G}^1$ resulting from a single edge addition to an acceptable graph at stage i; (b) rejects $\mathcal{G}^1$ if it conflicts (for all u) with $\mathcal{H}$; (c) otherwise keeps $\mathcal{G}^1$ as acceptable at $i+1$; and (d) if $\exists u[\mathcal{G}^u = \mathcal{H}]$, then adds $\mathcal{G}^1$ to $[\![\mathcal{H}]\!]$. RASL$_{ie}$ continues until there are no more edges to add (or it reaches stage n^2+1).

Results

Size of $[\![\mathcal{H}]\!]$ as a function of u for **100** random **5**-node SCCs with **25%** density, each of which was undersampled for $u \in \{2, \ldots, 11\}$.

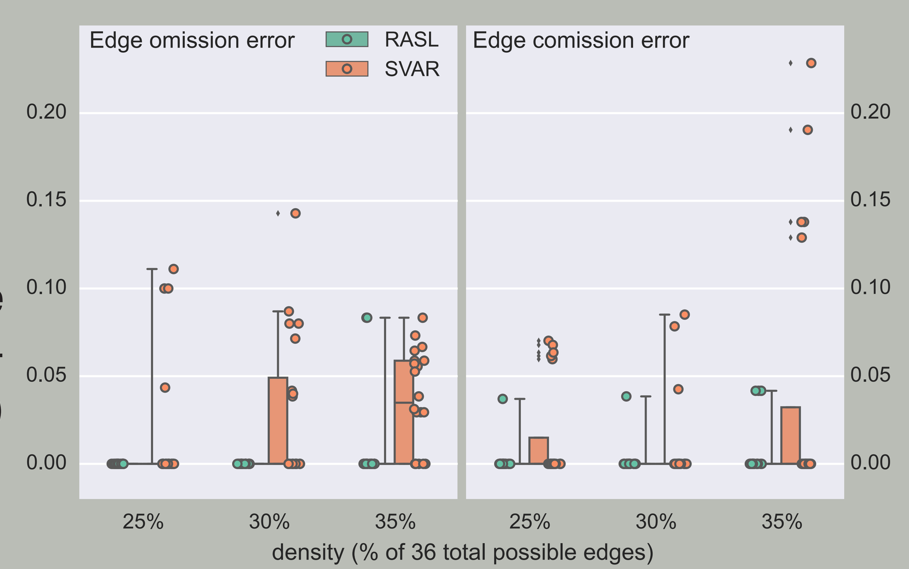


Size of $[\![\mathcal{H}]\!]$ as a function of density for **7**, **8** and **10**-node SCCs



Synthetic data $u = 2$

- generate a random SCC graph
- convert to a stable transition matrix
- simulate data via $\vec{x}_t = A\vec{x}_{t-1} + \vec{\eta}$
- direct minimization of the negative log likelihood of the structural vector autoregressive model (SVAR) $B\vec{x}_t = A\vec{x}_{t-1} + \vec{\eta}$



Conclusions

We here provided the first causal inference algorithms that can reliably learn causal structure from time series data when the system and measurement timescales diverge to an unknown degree. The RASL algorithms are complex, but not restricted to toy problems. We also showed that underdetermination of $\mathcal{G}^1$ is sometimes minimal, given the right methods. $[\![\mathcal{H}]\!]$ was often small; substantial system timescale causal structure could be learned from undersampled measurement timescale data. This paper has, however, expanded our causal inference "toolbox" to include cases of unknown undersampling.

http://www.cmu.edu NIH R01EB005846 NSF IIS-1318759 NSF IIS-1318815 & NIH U54HG008540 http://www.mrn.org