



# Extracting Stable Community Information on Relational Data

Kasimir Gabert<sup>1,2</sup>, Ali Pinar<sup>1</sup>

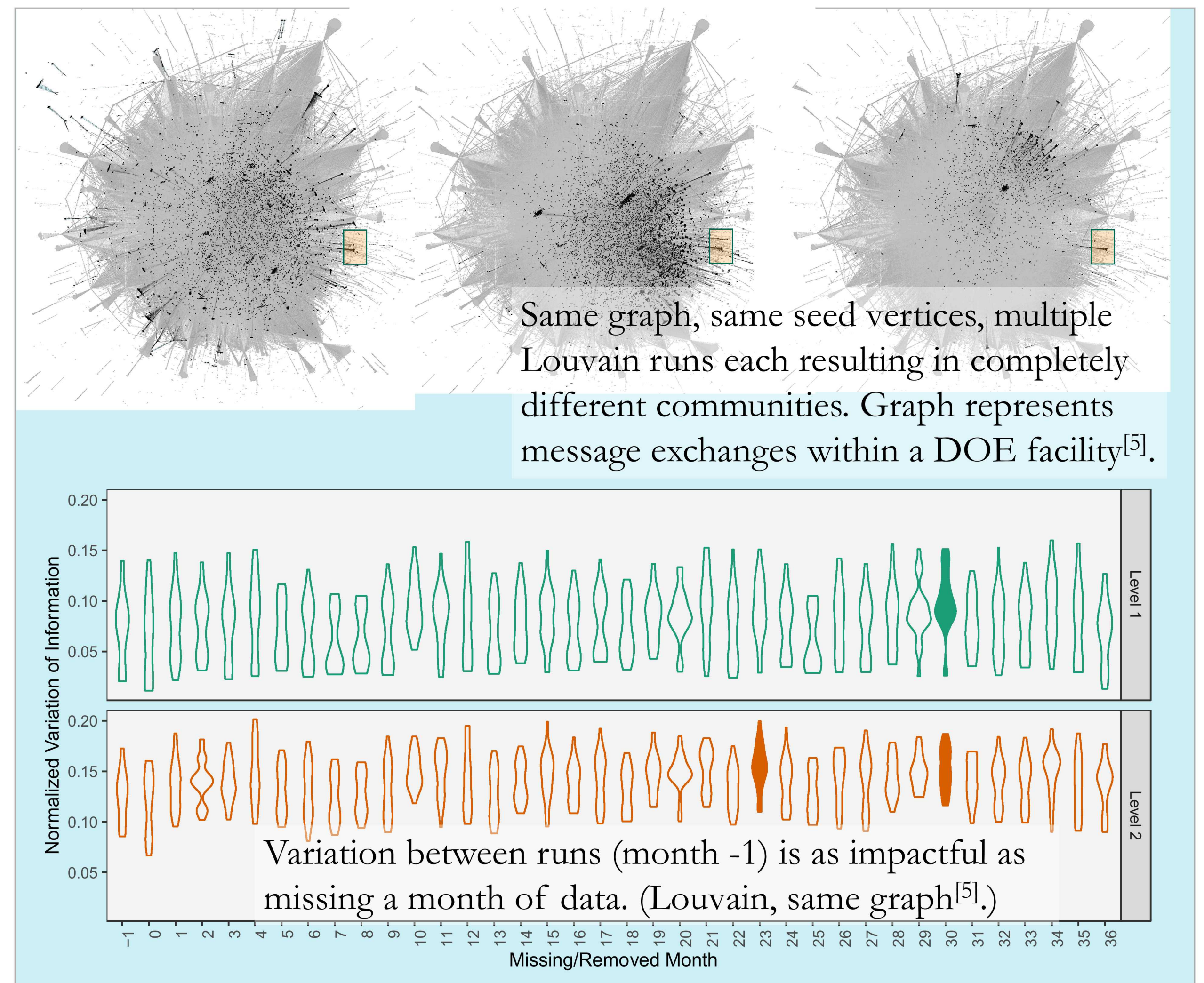
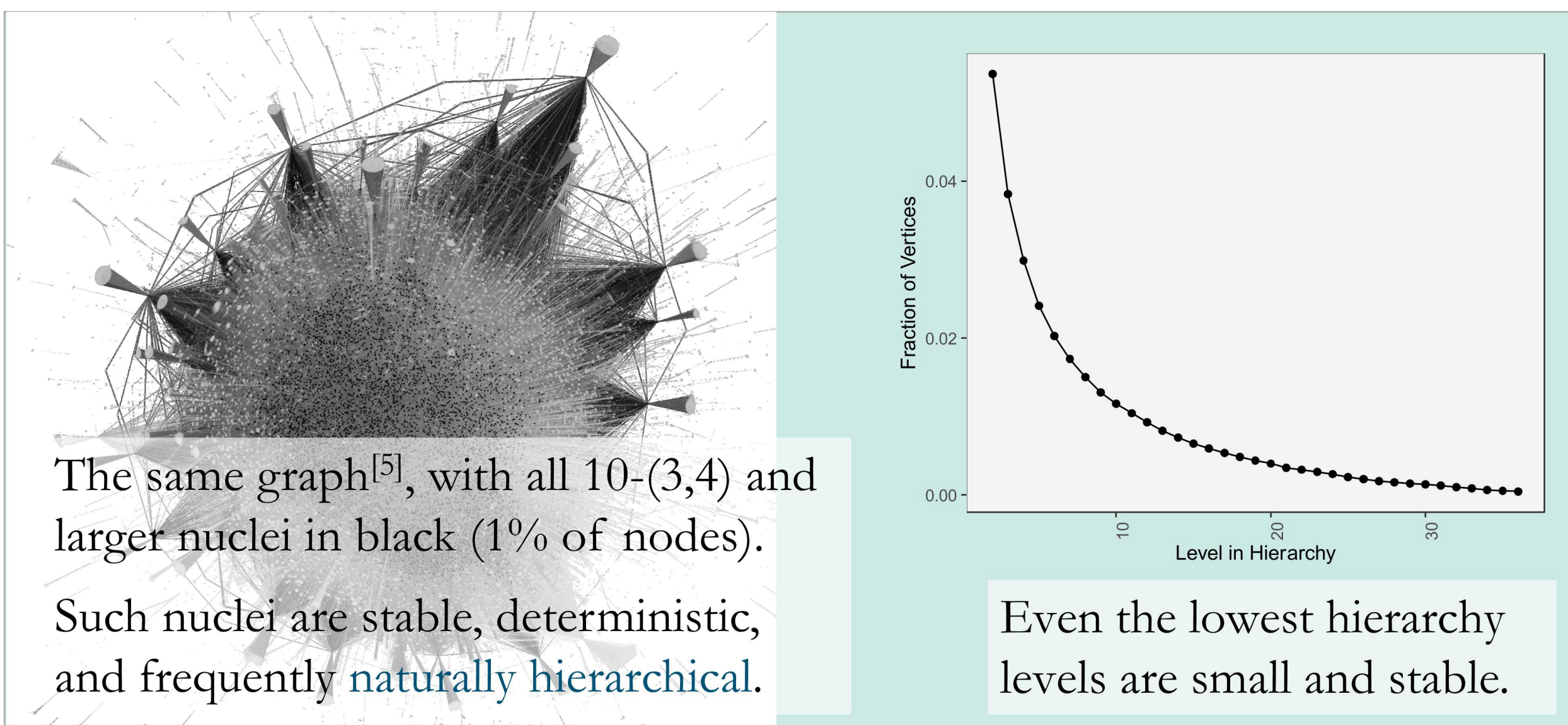
<sup>1</sup>Sandia National Laboratories, <sup>2</sup>Georgia Institute of Technology

## Problem:

- Community detection or clustering is highly desired across many national security datasets
- Popular algorithms may have wildly varying outputs on multiple runs or after minor changes to the data
- Can we trust outputs of such algorithms?

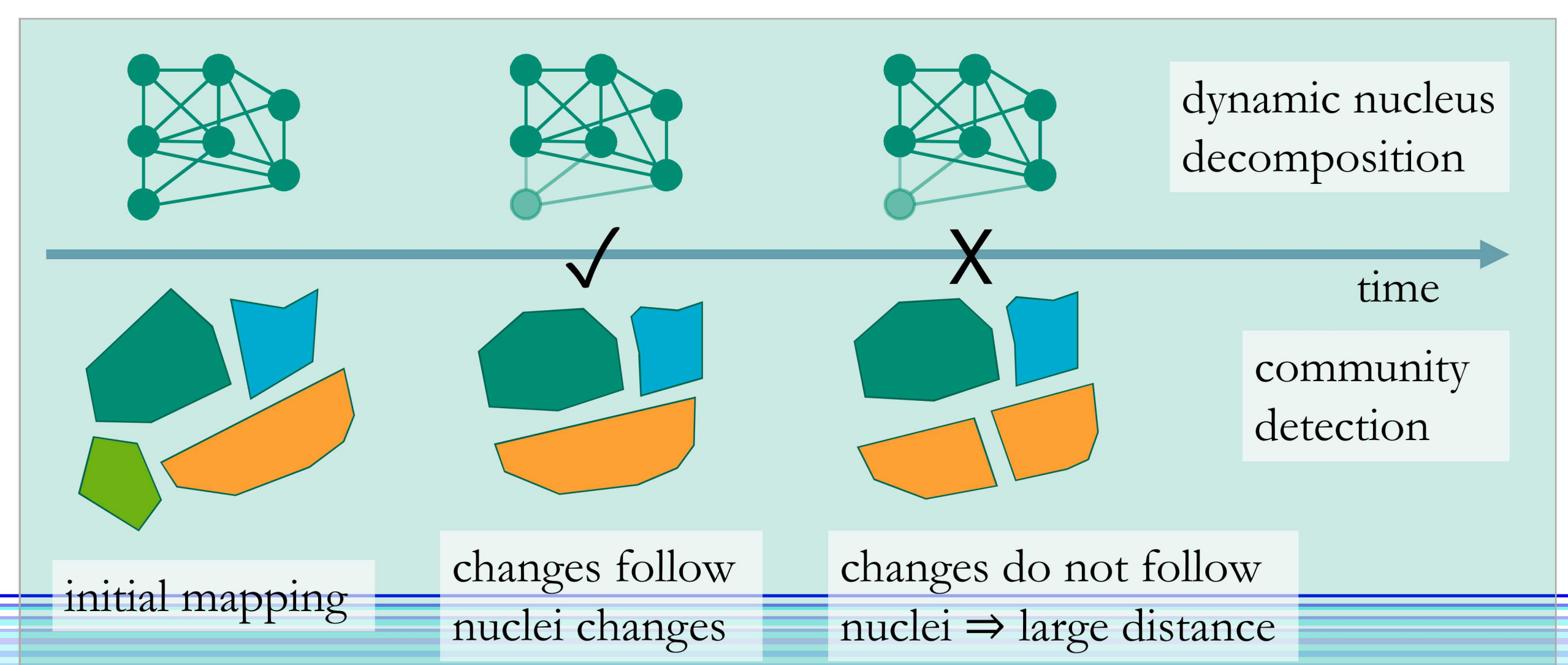
## Mitigation:

- Design stable algorithms that do not compromise solution quality
- Quantify distance between solutions
- Identify the smallest temporal/structural granularity that provides stability



## Dense Subgraphs:

- Graphs are globally sparse but locally dense
- There is a natural hierarchy among dense regions
- Nucleus decomposition finds useful and informative dense subgraphs efficiently
- We developed algorithms and implementations to track nuclei dynamically / temporally
- Our system is scalable, works on real datasets, and provides state-of-the-art solution quality



## Finding Stable Communities (on-going):

- Graph nuclei are deterministic and stable
- We connect communities to the nuclei within them
- Community changes can be tracked over time and temporal oscillations and abnormalities, with no corresponding nuclei changes, identified
- This technique can convert static into dynamic community detection and form a stability metric

## References:

1. Gabert, Pinar, Çatalyürek. "Computing Hierarchical Dense Structures on Dynamic Graph Streams." *In Submission*. 2020
2. Sariyüce, Seshadhri, Pinar, Çatalyürek. "Finding the hierarchy of dense subgraphs using nucleus decompositions." *WWW*. 2015.
3. Sariyüce, Pinar. "Fast Hierarchy Construction ...." *VLDB*. 2016.
4. Sariyüce, Seshadhri, Pinar. "Local algorithms for hierarchical dense subgraph discovery." *VLDB*. 2018.
5. PNNL Living Laboratory RPL Dataset

## Collaboration:

- Do you have a dataset where community detection is unstable or frustrating you? We are seeking additional applications and will be happy to collaborate
- Our software is available, please contact us for details

Contact: { kggaber, apinar }@sandia.gov