Scaling Causal AI for Seismic FEM Simulation
SCA/HPC Asia 2026

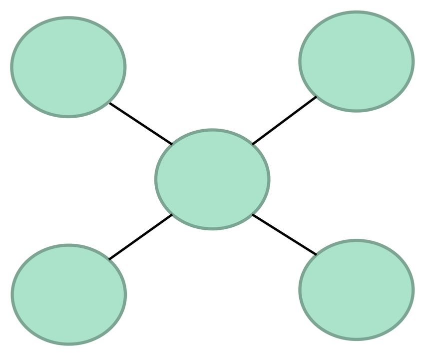
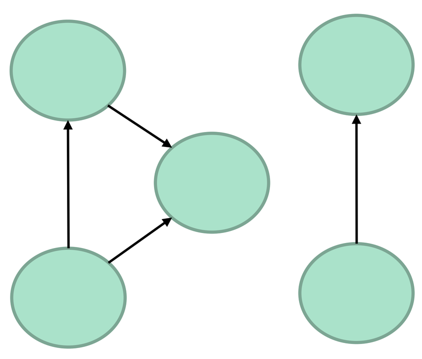
Poster n°142

Aurore Lomet, aurore.lomet@cea.fr
Mohd Afeef Badri, mohd-afeef.badri@cea.fr
Tristan Hascoet, tristan.hascoet@riken.jp

Causality for Seismic Data Assimilation

Beyond correlation-based data assimilation:

- For reliable risk assessment**, robust and large-scale high-performance computing of** earthquake **finite element method (FEM) simulations is needed.**
- The **robustness and predictive accuracy** of such simulations improve substantially when coupled with real-world observations through **data assimilation** techniques.
- Conventional data assimilation schemes often rely on **correlation-based coupling mechanisms** that can misrepresent the inherently **directional nature of physical interactions** in seismic wave propagation.

***Correlation modeling**: Undirected associations*

***Causal inference:** Directional cause-effect relations to reflect signal propagation*

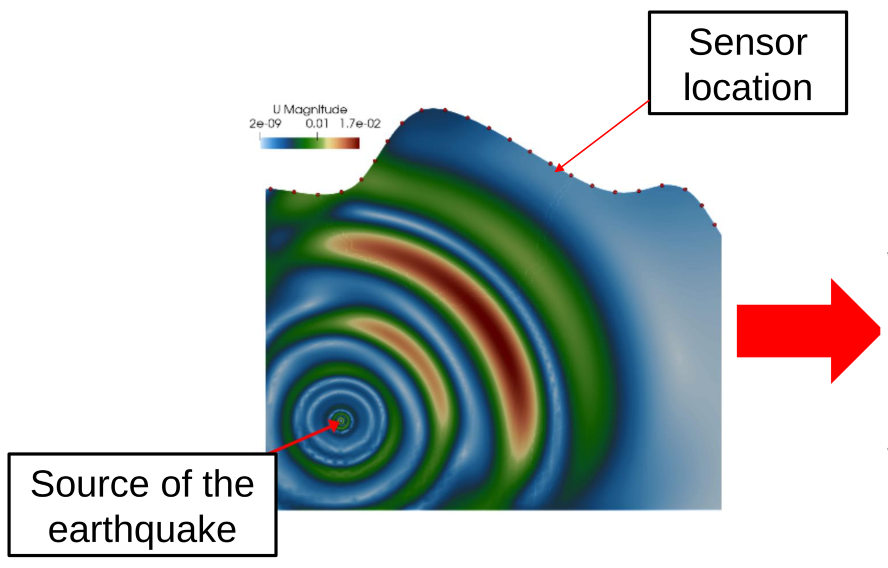
Fig 1. Correlation-based modeling vs directional cause–effect relations in seismic wave propagation.

Causal discovery for time series

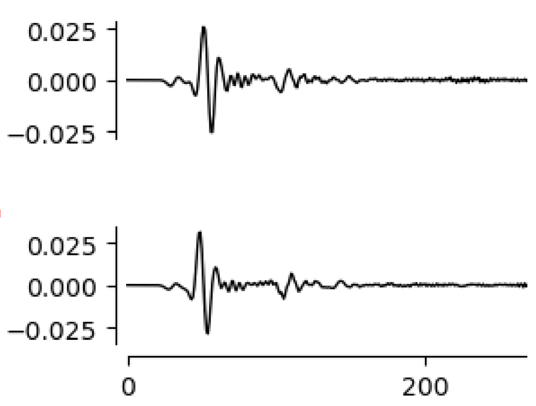
Causal inference identifies directional cause–effect relations among observational sensors

- To improve upon correlation based data assimilation, **causal inference methods** identify **directional cause-effect relations** among data assimilation sensors.
- Unlike correlation-based approaches, adapted **causal discovery** algorithms extract **temporal graphs** that track signal propagation and reveal internal structure with potential lags.
- For **time series data**, conditional independence tests determine whether past values of one signal predict future values of another, beyond the target's own history.
- With **minimal modeling assumptions**, this relies on nonparametric, kernel-based independence testing using the Hilbert–Schmidt Independence Criterion (HSIC).
- **Complexity of HSIC is $O(n^2)$–$O(n^3)$ for time and $O(n^2)$ for memory** due to $n \times n$ Gram matrices, limiting scalability on long time series which arises in earthquakes (with n is the length of each time series).

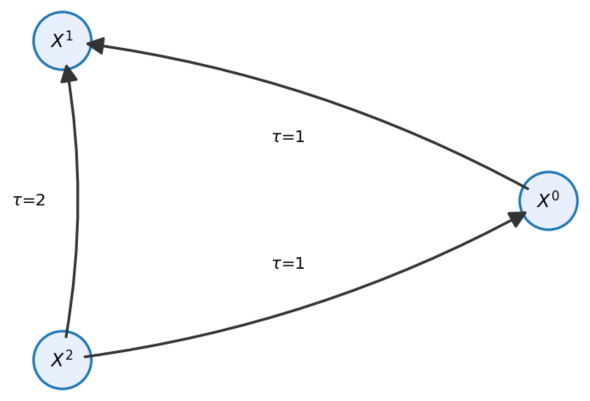
From earthquake simulation to temporal causal graphs



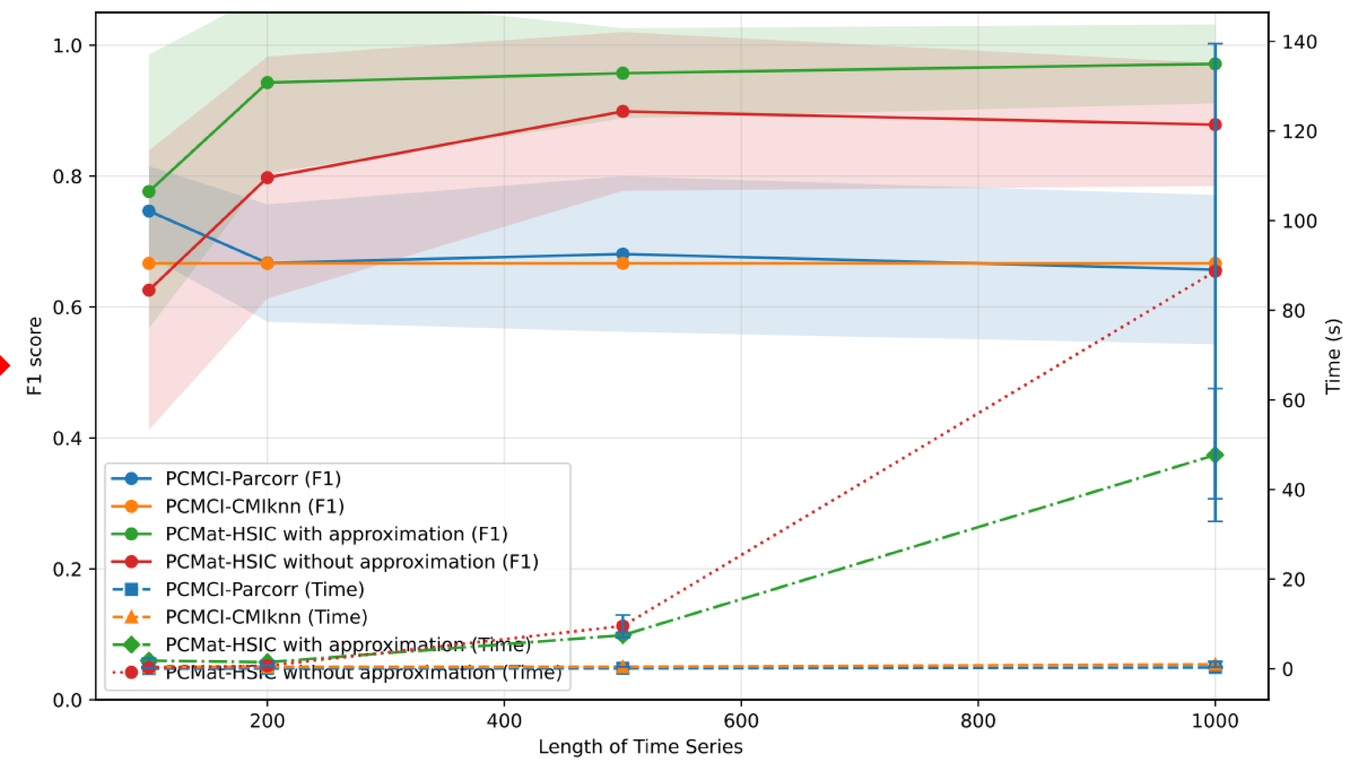
Earthquakes simulation based on FEM



Sensor measures following time



Causal graph learned from data considering temporal lag between sensors by constraint based model with HISC



F1-score and computation time across time-series lengths for PCMat-HSIC (with and without random Fourier features approximation) and for PCMCI-Parcorr and PCMCI-CMIknn

Fig 3. Causal graph learning and scaling with time-series length

Earthquake simulation based on the finite element method generates time series at sensor locations. A causal graph is learned from data by considering temporal lag between sensors using a constraint-based model with the Hilbert–Schmidt Independence Criterion. The right panel reports F1-score and computation time across time-series lengths for PCMat-HSIC (with and without random Fourier features approximation) and for PCMCI-Parcorr and PCMCI-CMIknn..

Algorithmic optimization

Low rank random Fourier features approximation:

- Towards delivering real-time causal inference over practically sized seismic sensor network, the present work proposes a new causal algorithm PCMat-HSIC and addresses its computational challenge through algorithmic and computational optimizations.
- On the algorithmic side, a **low rank random Fourier features approximation** of the kernel matrix by sampling from the Fourier spectrum reduces construction from $O(n^2)$–$O(n^3)$ to $O(nm)$ time and $O(nm)$ memory ($m \ll n$) with m the number of RFF

Porting the inference procedure to GPU

- The Hilbert-Schmidt Independence Criterion causal inference procedure is ported to **GPU** unit using the PyTorch library.
- Speedups up to 100 folds are reported on a NVIDIA A100 chip compared to a **CPU** baseline implementation.

A. Lomet[1], M-A. Badri[1], T. Hascoet[2]

1 Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation, F-91191, Gif-sur-Yvette, France
2 RIKEN RCCS & iTHEMS, Kobe, Japan

Experiments and results

- **Experiment**: Comparison of causality algorithms (PCMCI-Parcorr, PCMCI-CMIknn, PCMat-HSIC with random Fourier features approximation, PCMat-HSIC without approximation) using F1-score (score to recover the correct links of the causal graphs) and runtime across series lengths.
- **Results** (Figure): **PCMat-HSIC with random Fourier features approximation shows high F1-score with reduced runtime** relative to PCMat-HSIC without approximation as the length of time series increases; PCMCI-Parcorr and PCMCI-CMIknn show lower F1-score across the tested series lengths.

Conclusion and future work

- The present work proposes **a new causal algorithm PCMat-HSIC** and addresses its **computational challenge through algorithmic and computational optimizations**.
- **Future work**:
  - Integration of the random Fourier features approximation into the graphics processing unit port in order to deliver real-time inference over practically sized sensor networks
  - Comparison with other data assimilation methods.

Acknowledgments: Supported by the JHPCN project number jh230017 "New computational science through the fusion of computation, data, and learning in a heterogeneous environment". Project leaders: Kengo Nakajima; Evelyne Foerster.