Analysis Report: Convergent Cross Mapping (CCM) Performance on Simulated Causal Networks

Executive Summary

This report analyzes the performance of Convergent Cross Mapping (CCM) in detecting causality across three simulated network structures. The analysis reveals significant limitations of CCM in distinguishing between direct causality, indirect relationships through common drivers, and spurious correlations arising from deterministic coupling.

Background

Three simulation scenarios were designed to test CCM's ability to detect different types of causal relationships:

1. **Linear Chain**: Sequential causality (X1 → X2 → X3 → X4)
2. **Hub with Outsider**: X1 influences X2 (strong) and X3 (weak), X4 independent
3. **Complex Network**: X1 → X2 (strong), X1 → X3 (weak), X4 → X2 (weak)

Key Findings

1. Bidirectional Relationships in Unidirectional Systems (Simulation 1)

Observation: The supposedly unidirectional chain (X1 → X2 → X3 → X4) showed strong bidirectional relationships, with reverse causality (e.g., X3 → X1 = 0.993) being nearly as strong as forward causality.

Analysis: This unexpected bidirectionality arises from:

- **High coupling strength** (0.7) creating deterministic relationships where information can be inferred in both directions
- **Low noise levels** (0.1) allowing for precise reconstruction of past states
- **CCM's fundamental nature** of detecting predictability rather than true causal direction

Implication: In strongly coupled, low-noise systems, CCM may detect bidirectional relationships even when causality is strictly unidirectional.

2. Common Driver Misinterpretation (Simulation 2)

Observation: CCM detected a strong bidirectional relationship between X2 and X3 ($\rho = 0.999$), despite these variables having no direct causal connection.

Analysis: This finding illustrates a fundamental limitation of CCM when dealing with common drivers. When X1 independently influences both X2 and X3, the resulting correlation between X2 and X3 is misinterpreted by CCM as a direct causal relationship. This occurs because:

- Both variables become synchronized through their shared input from X1
- CCM detects the predictability between X2 and X3 without recognizing it stems from a common source
- The deterministic nature of the coupling (strength = 0.8) with low noise (0.1) exacerbates this effect

Implication: CCM cannot inherently distinguish between direct causality (A → B) and indirect correlation through a common driver (C → A, C → B).

3. Directional Asymmetry (Simulation 3)

Observation: The expected weak relationship X4 → X2 was not detected above the threshold ($\rho = 0.044$), while X2 → X4 showed a detectable relationship ($\rho = 0.476$).

Analysis: This asymmetry suggests that CCM's sensitivity to causal direction can be influenced by:

- The relative dynamics of the coupled variables
- The specific time delays and coupling strengths
- Potentially non-commutative effects of the CCM algorithm

Technical Limitations Identified

1. Parameter Sensitivity

The choice of coupling strengths (0.8 for strong, 0.3 for weak) may have been too high, leading to:

- Over-detection of relationships
- Reduced ability to distinguish between strong and weak causality
- Spurious bidirectional correlations

2. Noise-to-Signal Ratio

The low noise level (0.1) resulted in:

- Highly deterministic systems where past states are easily reconstructible
- Enhanced bidirectional relationships in unidirectional systems
- Reduced ability to distinguish genuine causal direction from mathematical correlation

3. Threshold Selection

The 0.3 threshold for significant relationships may be:

- Too low, leading to false positives
- Not appropriate for all relationship types
- Insufficient for distinguishing between direct and indirect causality

Recommendations for Improved Analysis

1. Parameter Optimization

- Reduce coupling strengths (e.g., strong = 0.5, weak = 0.1)
- Increase noise levels (e.g., 0.2) to introduce stochasticity
- Experiment with different threshold values (e.g., 0.5 for strong relationships)

2. Methodological Enhancements

- Implement **Conditional CCM** to control for common drivers
- Use **Partial Cross Mapping** to isolate direct relationships
- Consider **Transfer Entropy** as a complementary method for causal direction

3. System Design Modifications

- Introduce asymmetric coupling mechanisms
- Use different functional forms for different relationships (e.g., sin vs. cos)
- Implement variable time delays between different causal links

4. Validation Strategies

- Compare results with other causality detection methods
- Test on known ground-truth systems
- Perform sensitivity analysis on key parameters

Conclusions

This analysis demonstrates that while CCM is a powerful tool for detecting relationships in time series data, it has significant limitations in:

1. **Distinguishing causality types**: Direct vs. indirect, common driver effects
2. **Determining true causal direction**: Especially in strongly coupled, low-noise systems
3. **Avoiding false positives**: From deterministic correlations that are not causally meaningful

Users of CCM should be aware of these limitations and consider:

- The specific characteristics of their data (noise levels, coupling strengths)
- The interpretation of bidirectional relationships
- The possibility of common driver effects
- The need for complementary analysis methods

Future research should focus on developing enhanced versions of CCM that can better distinguish between different types of causal relationships and reduce the impact of system determinism on causal inference.

References

Note: This analysis is based on simulation results from three controlled scenarios designed to test CCM performance on different network structures. The findings are specific to the parameter settings and system designs used in these simulations.

Prepared by: Oded Kushnir, Ph.D. - Founding Director | Principal Data Scientist

Organization: Biodaat

Date: 15/05/2025

Supporting Data: See attached simulation logs (causality_results_20250515_161159/)