Causal Inference in GNM Proteins Using Gaussian Models: A Study of Residue Correlation and Transfer Entropy

bignozzi.1855163

September 2024

0.1 Background

The study of protein dynamics is crucial for understanding biological functions at the molecular level. The Gaussian Network Model (GNM) provides a simplified yet powerful framework for analyzing the collective motions of proteins based on their residue interactions. By modeling proteins as networks of nodes (residues) connected by springs (interactions), GNM allows researchers to predict fluctuations and dynamic behavior effectively.

Causal inference has emerged as a vital approach in biological systems, enabling researchers to discern relationships between variables beyond mere correlation. This thesis aims to integrate GNM with causal inference techniques to explore residue correlations and their implications on protein dynamics.

0.2 Objectives

The primary objectives of this research are:

- To analyze static and dynamic correlations between protein residues using a Gaussian model with cutoff.

- To compute response entropy and transfer entropy as measures of information flow within protein networks.

- To interpret the implications of these measures in understanding protein functionality.

0.3 Structure of the Thesis

This thesis is organized into six chapters. Chapter 2 reviews relevant literature on GNM, causal inference, and entropy measures. Chapter 3 details the methodology employed in this study, including data collection and analysis techniques. Chapter 4 presents the results obtained from the analysis. Chapter 5 discusses these results in the context of existing literature. Finally, Chapter 6 concludes the thesis and suggests directions for future research.

Chapter 1

Literature Review