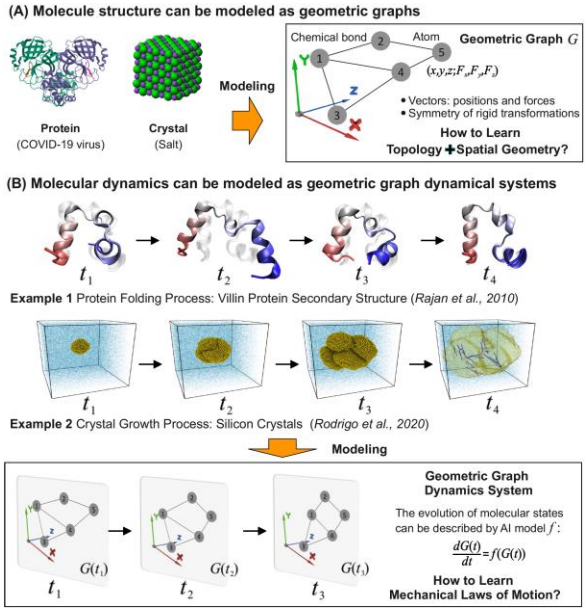
Toward Domain-Knowledge-Free Modeling: Bottleneck Analysis and Outlook for AI in Molecular Science

Hongbin PEI, Jingxin HAI, Zhewei WEI

Frontiers of Computer Science, DOI: 10.1007/s11704-026-51647-2

Problems

- Technical Bottlenecks of AI in Molecular Science:
  - Bottleneck of basis functions: Existing geometric basis functions suffer from incompatibility with large model architectures.
  - Bottleneck of dynamics models: Existing AI molecular dynamics models suffer from severe insufficient cross-molecular generalization.
  - Bottleneck of AI simulation systems: Existing AI-based molecular dynamics systems often fail in practical applications due to their inability to perform long-term simulations.



Geometric graph dynamical system learning is the fundamental AI problem for molecular structure and dynamics.

Ideas & Main Contributions

- Ideas: we advocate a domain-knowledge-free modeling paradigm for designing deep networks in AI4S. By this, we refer to models that do not rely on explicit domain knowledge modeling but instead employ simple, scalable, deeplearning-friendly architectures to learn scientific laws implicitly from large datasets, thereby enabling emergent intelligence in AI4S.

- Contributions:
  - The paper investigates limitations at three levels: geometric basis functions, dynamics models, and simulation systems.
  - We find that these issues are attributed to excessive explicit injection of domain knowledge that conflicts with deep learning scalability principles.
  - we advocate a domain-knowledge-free modeling paradigm for designing deep networks in AI4S, particularly for molecular learning.