

Wouter Boomsma

Department of Computer Science
University of Copenhagen

Title: Internal-Coordinate Density Modelling of Protein Structure: Covariance Matters

Abstract:

After the recent ground-breaking advances in protein structure prediction, one of remaining challenges in protein machine learning is to reliably predict distributions of structural states. Parametric models of small-scale fluctuations are difficult to fit due to complex covariance structures between degrees of freedom in the protein chain, often causing models to either violate local or global structural constraints. In this paper, we present a new strategy for modelling protein densities in internal coordinates, which uses constraints in 3D space to induce covariance structure between the internal degrees of freedom. We illustrate the potential of the procedure by constructing a variational autoencoder with full covariance output induced by the constraints implied by the conditional mean in 3D, and demonstrate that our approach makes it possible to scale density models of internal coordinates to full-size proteins.