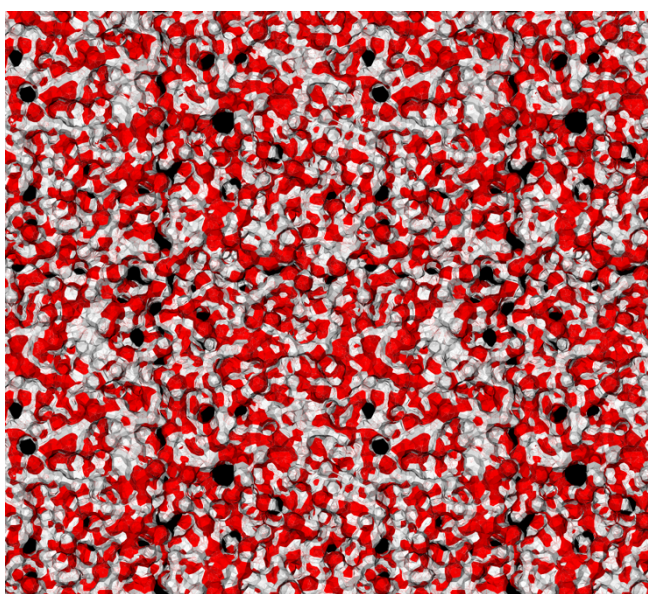


## Topological Data Analysis of Water Networks from Simulations

*Lee Steinberg & Jeremy Frey  
University of Southampton, TMCS  
ls2m15@soton.ac.uk*

Understanding the structure of water networks has major implications in the field of drug discovery. For example, understanding how the presence of a solute affects the native water structure could enable us to build better models for the prediction of free energy changes of solvation. Investigation of water networks has previously taken a graph-theoretical approach, defining a graph structure based on a measure of correlation between water molecules, such as in [1]. Such methods have also been used in studying radiation damage [2].



*Figure 1 - The surface of a simulated TIP3P water network at 300K. Holes are clearly visible in the system.*

However, these techniques require a user-input, for example defining a value of distance where two water molecules can be considered to be correlated. We propose an alternative approach based on the use of persistent homology [3], a technique in the field of topological data analysis (TDA), which defines a set of structures for the full range of parameter values, and considers how this series evolves to determine which features are the most useful.

We present an investigation into the structure of water networks in various models used in molecular dynamics simulations – TIP3P, TIP4P, and SPC. We discuss the adaptation of persistent homology methods to time-series data, such as the persistence landscape[4]. We relate persistent loops to the radial distribution function of the molecular dynamics model. We also discuss the presence of three-dimensional voids in water networks, and relate it to other properties.

- [1] B. Mooney, L. Corrales, A. Clark, *Molecular networks: An integrated graph theoretic and data mining tool to explore solvent organization in molecular simulation*, Journal of Computational Chemistry, 2012
- [2] A. Archer, H. Foxhall, N. Allen *et al*, *Order parameter and connectivity topology analysis of crystalline ceramics for nuclear waste immobilization*, Journal of Physics: Condensed Matter, 2014
- [3] H. Edelsbrunner, J. Harer, *Persistent Homology – A Survey*, Contemporary Mathematics, 2007
- [4] P. Bubenik, *Statistical Topological Data Analysis using Persistence Landscapes*, Journal of Machine Learning Research, 2015