ChemNetworks

A Complex Network Analysis Tool For Chemical Systems

User's Manual

Version 1.0

November 2013

Dr. Abdullah Ozkanlar and Prof. Aurora E. Clark

Department of Chemistry, Washington State University

Pullman, Washington 99164

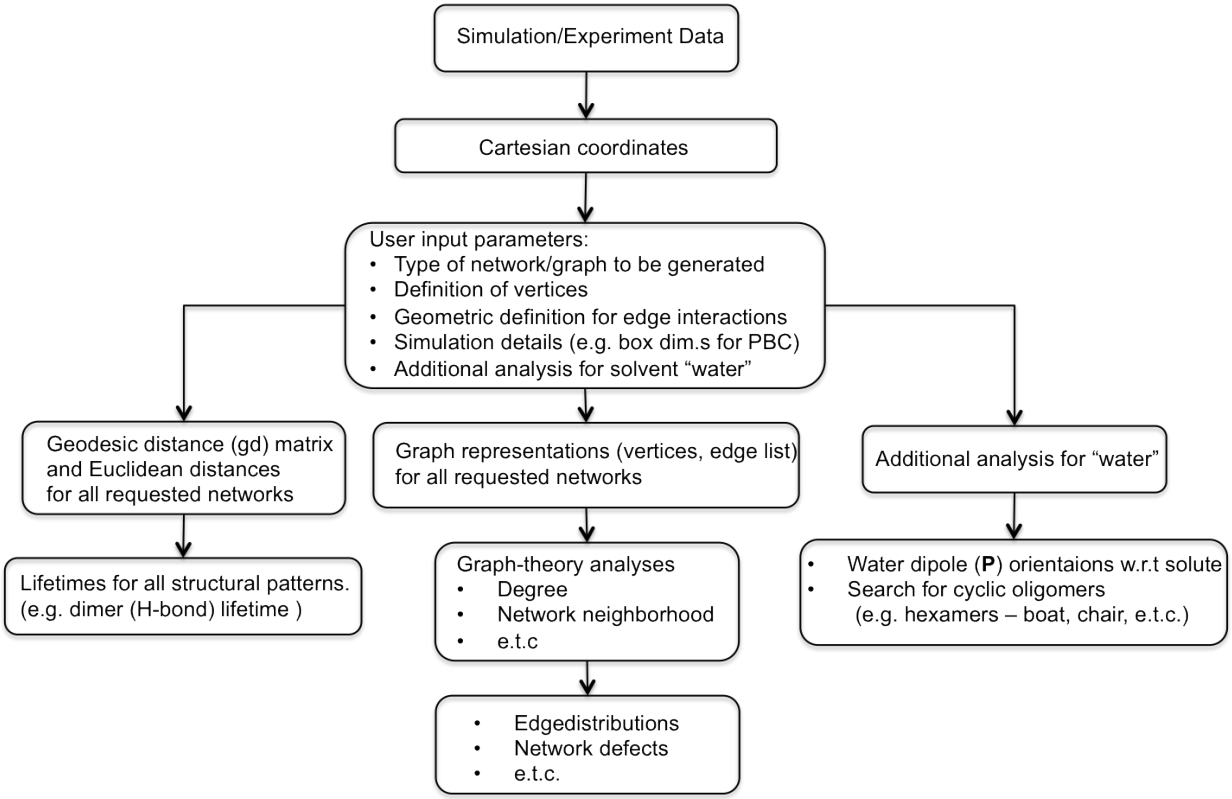
TABLE OF CONTENTS

ChemNetworks Manual

1.0 INTRODUCTION

ChemNetworks is a multipurpose tool for performing topological analyses of chemical networks defined by intermolecular interactions. It processes the Cartesian coordinates of chemical systems into network/graph formalism and applies topological network analyses including the Network neighborhood, the determination of geodesic pathways, the degree census, direct structural searches, and the distribution of defect states of network. These properties can help to understand the network patterns and organization that may influence physical properties and chemical reactivity. *ChemNetworks* is a series of codes written primarily in C, with some limited post-processing performed using scripts in the R statistical computing language. This software is a completely generalized tool that can be used to understand a very large range of chemical systems that include complex solutions, liquid interfaces, self-assemblies, or pure liquids undergoing phase changes.

Workflow for *ChemNetworks*



2.0 INSTALLATION AND COMPILATION

The *ChemNetworks* package is freeware and can be downloaded from the website aclark.chem.wsu.edu/software. It consists of the source files written in C, a header file, a make file for compilation, example input and output files stored in the Examples folder, utilities folder including a script written in R language for post-processing, and this user's manual.

In order to compile *ChemNetworks* source code, use the make file provided:

make -f MakeFile

This will generate an executable named ChemNetworks.exe. This executable reads the input file and the .xyz Cartesian coordinate files for the chemical species under consideration.

3.0 RUNNING *ChemNetworks*

Use following syntax to run *ChemNetworks*:

./ChemNetworks.exe {Inputfile.input} {Solvent.xyz files} {Solute.xyz files}

For example, for a chemical system of a ternary solution phase mixture, use following;

./ChemNetworks.exe Inputfile Solvent1 Solvent2 Solvent3

Inputfile: the input file for *ChemNetworks*

Solvent1: the .xzy filename for the 1st solvent defined in the input file

Solvent2: the .xzy filename for the 2nd solvent defined in the input file

Solvent3: the .xzy filename for the 3rd solvent defined in the input file

Note that the order of the filenames used when running the executable and the order of the chemical species defined in the input file must match.

4.0 KEYWORD LIST

ChemNetworks is designed to work primarily with chemical mixtures and thus the notation of a major component "solvent" and a minor component "solute" is adopted. The list of all keywords is given in this section. The minimum required keywords are highlighted with **bold** fonts. The use of the other keywords depends on the user input for the required keywords. The values entered for the required keywords determine the additional keywords that are needed. Therefore, the input file is highly dynamic and specific for the type of chemical system of interest and type of the desired network and network analyses. See Examples (Section 6) for sample input files.

[NUMBER OF SOLVENT TYPES] The number of different types of "solvent" species. Max. is 3.

[NUMBER OF SOLUTE TYPES] The number of different types of "solute" species. Max. is 2.

[NUMBER OF ATOMS IN SOLVENT1] The number of atoms in "solvent1" molecule. Then, the atom labels with their positions/orders within the molecule must be written. These must be consistent with the corresponding .xyz file.

[NUMBER OF ATOMS IN SOLVENT2] Must be specified if the [NUMBER OF SOLVENT TYPES] has value of 2.

[NUMBER OF ATOMS IN SOLVENT3] Must be specified if the [NUMBER OF SOLVENT TYPES] has value of 3.

[NUMBER OF ATOMS IN SOLUTE1] Must be specified if the [NUMBER OF SOLUTE TYPES] has value of 1. The number of atoms in "solute1" molecule. Then, the atom labels with their positions in the molecule must be written. These must be consistent with the corresponding .xyz file.

[NUMBER OF ATOMS IN SOLUTE2] Must be specified if the [NUMBER OF SOLUTE TYPES] has value of 2.

[PERIODIC BOUNDARY CONDITIONS] 1 (Yes, take into account the periodic boundary conditions); 0 (No)

[BOX XSIDE] Specify the rectangular simulation box dimensions (x-side) if the value of [PERIODIC BOUNDARY CONDITIONS] is 1.

[BOX YSIDE] The simulation box dimensions (y-side)

[BOX ZSIDE] The simulation box dimensions (z-side)

[GRAPH SOLVENT1 SOLVENT1] 1 (Yes, construct the H-bond network of "solvent1"); 0 (No)

[SOLVENT1 SOLVENT1 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT1] is 1. The H-bond distance criteria between "solvent1" molecules are specified between the pairs of atoms of "solvent1" molecules.

[SOLVENT1 SOLVENT1 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT1] is 1. The H-bond angle criteria between "solvent1" molecules are specified between the atoms of "solvent1" molecules.

[GRAPH SOLVENT2 SOLVENT2] 1 (Yes, construct the H-bond network of "solvent2"); 0 (No)

[SOLVENT1 SOLVENT2 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT2 SOLVENT2] is 1. The H-bond distance criteria between "solvent2" molecules are specified between the pairs of atoms of "solvent2" molecules.

[SOLVENT1 SOLVENT2 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT2 SOLVENT2] is 1. The H-bond angle criteria between "solvent2" molecules are specified between the atoms of "solvent2" molecules.

[GRAPH SOLVENT3 SOLVENT3] 1 (Yes, construct the H-bond network of "solvent3"); 0 (No)

[SOLVENT3 SOLVENT3 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT3 SOLVENT3] is 1. The H-bond distance criteria between "solvent3" molecules are specified between the pairs of atoms of "solvent3" molecules.

[SOLVENT3 SOLVENT3 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT3 SOLVENT3] is 1. The H-bond angle criteria between "solvent3" molecules are specified between the atoms of "solvent3" molecules.

[GRAPH SOLVENT1 SOLVENT2] 1 (Yes, construct the network formed between "solvent1" and "solvent2" H-bonds); 0 (No)

[SOLVENT1 SOLVENT2 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT2] is 1. The H-bond distance criteria between "solvent1" and "solvent2" molecules are specified between the pairs of atoms of "solvent1" and "solvent2" molecules.

[SOLVENT1 SOLVENT2 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT2] is 1. The H-bond angle criteria between "solvent1" and "solvent2"molecules are specified between the atoms of "solvent1" and "solvent2" molecules.

[GRAPH SOLVENT1 SOLVENT3] 1 (Yes, construct the network formed between "solvent1" and "solvent3" H-bonds); 0 (No)

[SOLVENT1 SOLVENT3 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT3] is 1. The H-bond distance criteria between "solvent1" and "solvent3" molecules are specified between the pairs of atoms of "solvent1" and "solvent3" molecules.

[SOLVENT1 SOLVENT3 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT1 SOLVENT3] is 1. The H-bond angle criteria between "solvent1" and "solvent3" molecules are specified between the atoms of "solvent1" and "solvent3" molecules.

[GRAPH SOLVENT1 SOLUTE1] 1 (Yes, construct the network formed between "solvent1" and "solute1" H-bonds); 0 (No)

[SOLVENT1 SOLUTE1 CUTOFF] Must be specified if the value of [GRAPH SOLVENT1 SOLUTE1] is 1. The cutoff distance criteria between "solvent1" and "solute1" molecules are specified between the pairs of atoms of "solvent1" and "solute1" molecules.

[GRAPH SOLVENT1 SOLUTE2] 1 (Yes, construct the network formed between "solvent1" and "solute2" H-bonds); 0 (No)

[SOLVENT1 SOLUTE2 CUTOFF] Must be specified if the value of [GRAPH SOLVENT1 SOLUTE2] is 1. The cutoff distance criteria between "solvent1" and "solute2" molecules are specified between the pairs of atoms of "solvent1" and "solute2" molecules.

[GRAPH SOLVENT2 SOLVENT3] 1 (Yes, construct the network formed between "solvent2" and "solvent3" H-bonds); 0 (No)

[SOLVENT2 SOLVENT3 HBOND DISTANCE] Must be specified if the value of [GRAPH SOLVENT2 SOLVENT3] is 1. The H-bond distance criteria between "solvent2" and "solvent3" molecules are specified between the pairs of atoms of "solvent2" and "solvent3" molecules.

[SOLVENT2 SOLVENT3 HBOND ANGLE] Must be specified if the value of [GRAPH SOLVENT2 SOLVENT3] is 1. The H-bond angle criteria between "solvent2" and "solvent3" molecules are specified between the atoms of "solvent2" and "solvent3" molecules.

[GRAPH SOLVENT2 SOLUTE1] 1 (Yes, construct the network formed between "solvent2" and "solute1" H-bonds); 0 (No)

[SOLVENT2 SOLUTE1 CUTOFF] Must be specified if the value of [GRAPH SOLVENT2 SOLUTE1] is 1. The cutoff distance criteria between "solvent2" and "solute1" molecules are specified between the pairs of atoms of "solvent2" and "solute1" molecules.

[GRAPH SOLVENT2 SOLUTE2] 1 (Yes, construct the network formed between "solvent2" and "solute2" H-bonds); 0 (No)

[SOLVENT2 SOLUTE2 CUTOFF] Must be specified if the value of [GRAPH SOLVENT2 SOLUTE2] is 1. The cutoff distance criteria between "solvent2" and "solute2" molecules are specified between the pairs of atoms of "solvent2" and "solute2" molecules.

[GRAPH SOLVENT3 SOLUTE1] 1 (Yes, construct the network formed between "solvent3" and "solute1" H-bonds); 0 (No)

[SOLVENT3 SOLUTE1 CUTOFF] Must be specified if the value of [GRAPH SOLVENT3 SOLUTE1] is 1. The cutoff distance criteria between "solvent3" and "solute1" molecules are specified between the pairs of atoms of "solvent3" and "solute1" molecules.

[GRAPH SOLVENT3 SOLUTE2] 1 (Yes, construct the network formed between "solvent3" and "solute2" H-bonds); 0 (No)

[SOLVENT3 SOLUTE2 CUTOFF] Must be specified if the value of [GRAPH SOLVENT3 SOLUTE2] is 1. The cutoff distance criteria between "solvent3" and "solute2" molecules are specified between the pairs of atoms of "solvent3" and "solute2" molecules.

[GRAPH SOLVENT1 SOLVENT2 SOLVENT3] 1 (Yes, construct the H-bond network of entire system of "solvent1", "solvent2" and "solvent3"); 0 (No)

[PRINT NUMBER OF NODES] 1 (Yes, print the number of nodes within every graph); 0 (No)

[SOLUTE1 WATER DIPOLE ORIENTATIONS] 1 N ← 1: (Yes, get the water dipole angles w.r.t. the "solute1" species); **N**: Water is the N^{th} "solvent" defined in the input file.

[SOLUTE2 WATER DIPOLE ORIENTATIONS] 0 (No)

[SOLVENT WATER DIPOLE ORIENTATIONS] 1 N M ← 1: (Yes, get the water dipole angles w.r.t. the M^{th} "solvent" defined in the input file); **N**: Water is the N^{th} "solvent" defined in the input file.

[GEODESICS GD] 1 (Yes, calculate the geodesic distance (gd) matrix and Euclidean distance for all requested graph/network type. Then, all keywords below starting with GD are required); 0 (No)

[GD SOLVENT1] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT1 SOLVENT1] keyword); 0 (No)

[GD SOLVENT1 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT1 SOLVENT1] keyword); 0 (No)

[GD SOLVENT1 EUCLIDEAN REFERENCE] N ← Calculate the Euclidean distance between the reference atoms, whose order number is N as defined in [NUMBER OF ATOMS IN SOLVENT1] keyword, of terminal nodes.

[GD SOLVENT2] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT2 SOLVENT2] keyword); 0 (No)

[GD SOLVENT2 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT2 SOLVENT2] keyword); 0 (No)

[GD SOLVENT2 EUCLIDEAN REFERENCE] N ← Calculate the Euclidean distance between the reference atoms, whose order number is N as defined in [NUMBER OF ATOMS IN SOLVENT2] keyword, of terminal nodes.

[GD SOLVENT3] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT3 SOLVENT3] keyword); 0 (No)

[GD SOLVENT3 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT3 SOLVENT3] keyword); 0 (No)

[GD SOLVENT3 EUCLIDEAN REFERENCE] N ← Calculate the Euclidean distance between the reference atoms, whose order number is N as defined in [NUMBER OF ATOMS IN SOLVENT3] keyword, of terminal nodes.

[GD SOLVENT1 SOLVENT2] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT1 SOLVENT2] keyword); 0 (No)

[GD SOLVENT1 SOLVENT2 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT1 SOLVENT2] keyword); 0 (No)

[GD SOLVENT1 SOLVENT2 EUCLIDEAN REFERENCE] M N ← Calculate the Euclidean distance between the reference atoms, whose order number are M for “solvent1” and N for “solvent2”, of terminal nodes.

[GD SOLVENT1 SOLVENT3] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT1 SOLVENT3] keyword); 0 (No)

[GD SOLVENT1 SOLVENT3 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT1 SOLVENT3] keyword); 0 (No)

[GD SOLVENT1 SOLVENT3 EUCLIDEAN REFERENCE] M N ← Calculate the Euclidean distance between the reference atoms, whose order number are M for “solvent1” and N for “solvent3”, of terminal nodes.

[GD SOLVENT2 SOLVENT3] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT2 SOLVENT3] keyword); 0 (No)

[GD SOLVENT2 SOLVENT3 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT2 SOLVENT3] keyword); 0 (No)

[GD SOLVENT2 SOLVENT3 EUCLIDEAN REFERENCE] M N ← Calculate the Euclidean distance between the reference atoms, whose order number are M for "solvent2" and N for "solvent3", of terminal nodes.

[GD SOLVENT1 SOLVENT2 SOLVENT3] 1 (Yes, calculate gd matrix for the network obtained from [GRAPH SOLVENT1 SOLVENT2 SOLVENT3] keyword); 0 (No)

[GD SOLVENT1 SOLVENT2 SOLVENT3 EUCLIDEAN DISTANCE] 1 (Yes, calculate Euclidean distance between the terminal nodes of all geodesic pathways within the network obtained from [GRAPH SOLVENT1 SOLVENT2 SOLVENT3] keyword); 0 (No)

[GD SOLVENT1 SOLVENT2 SOLVENT3 EUCLIDEAN REFERENCE] M N P← Calculate the Euclidean distance between the reference atoms (whose order number are M for "solvent1", N for "solvent2", and P for "solvent3") of terminal nodes.

[WATER STRUCTURES] 1 N ← 1: (Yes, search for water oligomers. Then, all keywords below starting with WATER are required), 0 (No); **N**: Water is the Nth "solvent" defined in the input file.

[WATER HEXAMER RING] 1 (Yes, search for cyclic ring hexamers); 0 (No)

[WATER HEXAMER BOOK] 1 (Yes, search for book hexamers); 0 (No)

[WATER HEXAMER PRISM] 1 (Yes, search for prism hexamers); 0 (No)

[WATER HEXAMER CAGE] 1 (Yes, search for cage hexamers); 0 (No)

[WATER HEXAMER BAG] 1 (Yes, search for bag hexamers); 0 (No)

[WATER HEXAMER BOAT] 1 (Yes, search for boat hexamers); 0 (No)

[WATER HEXAMER CHAIR] 1 (Yes, search for chair hexamers); 0 (No)

[WATER HEXAMER PRISMBOOK] 1 (Yes, search for prismbook hexamers); 0 (No)

[WATER PENTAMER SEARCH] 1 (Yes, search for cyclic pentamers); 0 (No)

[WATER TETRAMER SEARCH] 1 (Yes, search for cyclic tetramers); 0 (No)

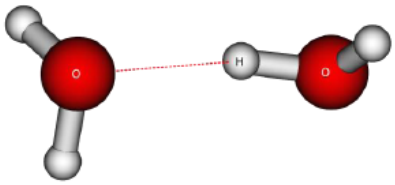
[WATER TRIMER SEARCH] 1 (Yes, search for cyclic trimers); 0 (No)

[WATER ISOLATED STRUCTURES] 1 (Yes, search for all isolated oligomers requested above); 0 (No)

5.0 GEOMETRIC CRITERIA FOR WATER OLIGOMERS

Direct structural search of specific water oligomers is invoked using the [WATER STRUCTURES] keyword as introduced in previous section. There has been historically a strong interest in the presence of polygons and cyclic oligomers of the form $(H_2O)_{3\text{-}5}$, and water hexamers, $(H_2O)_6$ (the bag, boat, book, cage, chair, prism, prismbook, and ring configurations). *ChemNetworks* incorporates geometric criteria for vertex angles and dihedrals for these structures, allowing for thermally induced angular deformations. The details are given below.

Dimer

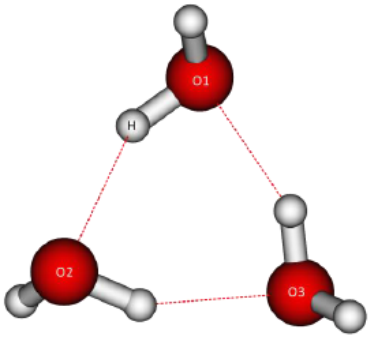


Definition of H-bonding:

$(\angle O \cdots H - O) > \theta$ (e.g. $\theta = 150°$)

$d(O \cdots H) < d$ (e.g. $d = 2.5$Å)

Cyclic Trimer

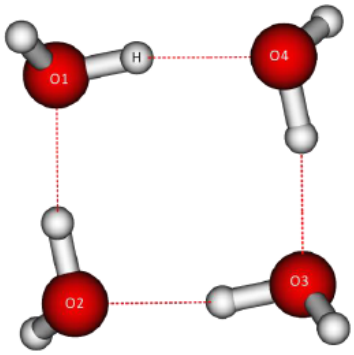


$50° < (\angle O1 - O2 - O3) < 70°$

$50° < (\angle O2 - O3 - O1) < 70°$

$50° < (\angle O3 - O1 - O2) < 70°$

Cyclic Tetramer (square)



$80° < (\angle O1 - O2 - O3) < 100°$

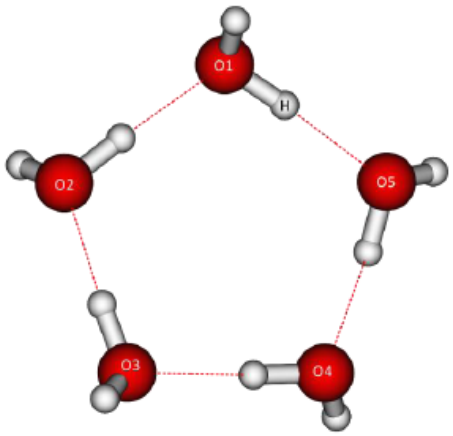
$80° < (\angle O2 - O3 - O4) < 100°$

$80° < (\angle O3 - O4 - O1) < 100°$

$80° < (\angle O4 - O1 - O2) < 100°$

$|\angle O1 - O2 - O3 - O4| \leq 20°$

Cyclic Pentamer



$$98° < (\angle O1 - O2 - O3) < 118°$$

$$98° < (\angle O2 - O3 - O4) < 118°$$

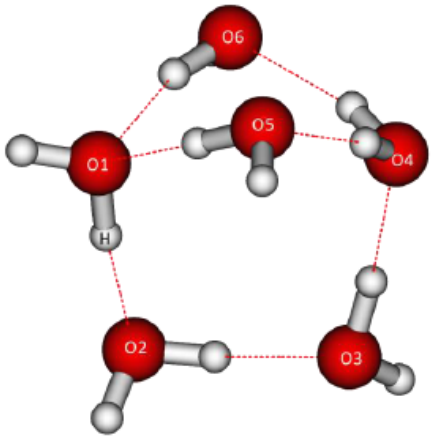
$$98° < (\angle O3 - O4 - O5) < 118°$$

$$98° < (\angle O4 - O5 - O1) < 118°$$

$$|\angle O1 - O2 - O3 - O4| \leq 20°$$

$$|\angle O1 - O5 - O4 - O3| \leq 20°$$

Hexamer Bag



$$80° \leq (\angle O1 - O2 - O3) \leq 140°$$

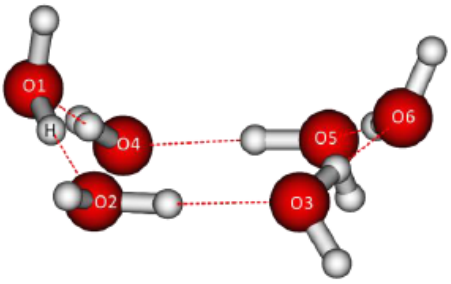
$$80° \leq (\angle O2 - O3 - O4) \leq 140°$$

$$0° \leq |\angle O1 - O2 - O3 - O4| \leq 25°$$

$$60° \leq |\angle O5 - O1 - O4 - O6| \leq 180°$$

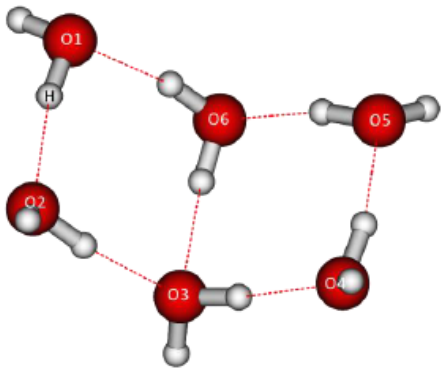
$$60° \leq |\angle O6 - O4 - O1 - O5| \leq 180°$$

Hexamer Boat



$$100° \leq (\angle O1 - O2 - O3) \leq 130°$$
$$100° \leq (\angle O1 - O4 - O5) \leq 130°$$
$$100° \leq (\angle O6 - O5 - O4) \leq 130°$$
$$100° \leq (\angle O6 - O3 - O2) \leq 130°$$
$$80° \leq (\angle O3 - O2 - O4) \leq 100°$$
$$80° \leq (\angle O2 - O4 - O5) \leq 100°$$
$$|\angle O2 - O3 - O5 - O4| \leq 15°$$
$$10° \leq |\angle O3 - O4 - O2 - O1| \leq 90°$$
$$10° \leq |\angle O4 - O3 - O5 - O6| \leq 90°$$

Hexamer Book



$80° < (\angle O1 - O2 - O3) < 100°$

$80° < (\angle O2 - O3 - O4) < 160°$

$80° < (\angle O3 - O4 - O5) < 100°$

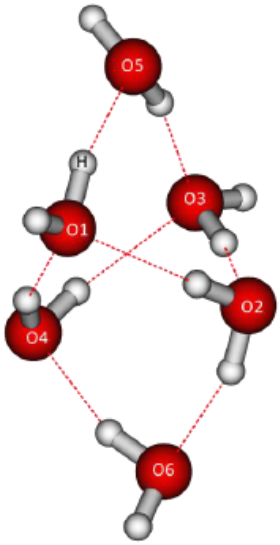
$80° < (\angle O4 - O5 - O6) < 100°$

$80° < (\angle O5 - O6 - O1) < 160°$

$|\angle O1 - O2 - O3 - O6| \leq 20°$

$|\angle O3 - O4 - O5 - O6| \leq 20°$

Hexamer Cage



$(\angle O \cdots H - O) > 130°$[a], for all waters

$60° < (\angle O1 - O2 - O3) < 120°$

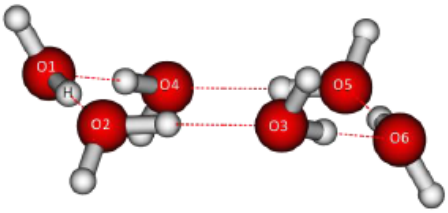
$60° < (\angle O2 - O3 - O4) < 120°$

$40° < |\angle O1 - O2 - O3 - O4| < 90°$

$150° < |\angle O5 - O1 - O3 - O6| \leq 180°$

$150° < |\angle O6 - O2 - O4 - O5| \leq 180°$

Hexamer Chair



$100° \leq (\angle O1 - O2 - O3) \leq 130°$

$100° \leq (\angle O1 - O4 - O5) \leq 130°$

$100° \leq (\angle O6 - O5 - O4) \leq 130°$

$100° \leq (\angle O6 - O3 - O2) \leq 130°$

$80° \leq (\angle O3 - O2 - O4) \leq 100°$

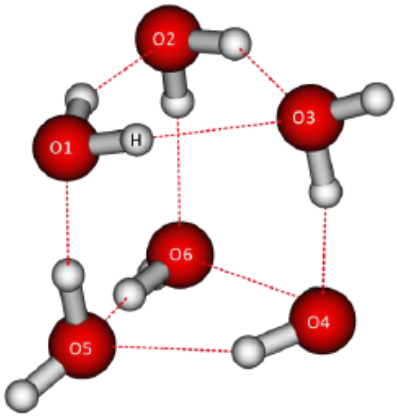
$80° \leq (\angle O2 - O4 - O5) \leq 100°$

$|\angle O2 - O3 - O5 - O4| \leq 15°$

$10° \leq |\angle O3 - O4 - O2 - O1| \leq 90°$

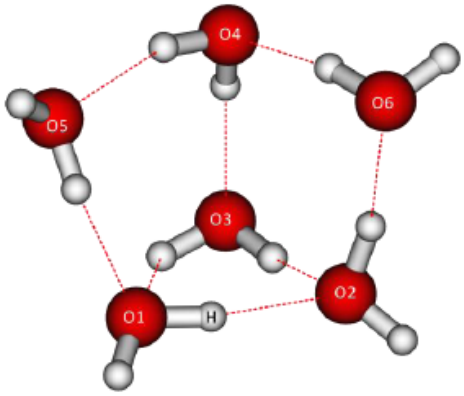
$10° \leq |\angle O4 - O3 - O5 - O6| \leq 90°$

Hexamer Prism



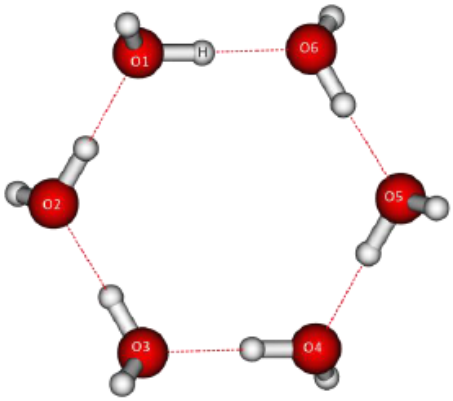
$50° < (\angle O1 - O2 - O3) < 70°$
$50° < (\angle O4 - O5 - O6) < 70°$
$80° < (\angle O2 - O3 - O4) < 100°$
$80° < (\angle O3 - O4 - O5) < 100°$
$80° < (\angle O5 - O6 - O2) < 100°$
$(\angle O2 \cdots H - O1) > 120°^{a}$
$(\angle O3 \cdots H - O2) > 120°^{a}$
$(\angle O3 \cdots H - O1) > 120°^{a}$
$(\angle O5 \cdots H - O4) > 120°^{a}$
$(\angle O6 \cdots H - O4) > 120°^{a}$
$(\angle O5 \cdots H - O6) > 120°^{a}$
$(\angle O4 \cdots H - O3) > 140°^{a}$
$(\angle O1 \cdots H - O5) > 140°^{a}$
$(\angle O6 \cdots H - O2) > 140°^{a}$

Hexamer Prism/Book



$50° < (\angle O1 - O2 - O3) < 70°$
$80° < (\angle O2 - O3 - O4) < 100°$
$80° < (\angle O3 - O4 - O5) < 100°$
$(\angle O5 - O4 - O6) > 70°$
$(\angle O2 \cdots H - O1) > 120°^{a}$
$(\angle O2 \cdots H - O3) > 120°^{a}$
$(\angle O1 \cdots H - O3) > 120°^{a}$
$(\angle O5 \cdots H - O4) > 120°^{a}$
$(\angle O4 \cdots H - O6) > 120°^{a}$
$(\angle O3 \cdots H - O4) > 140°^{a}$
$(\angle O1 \cdots H - O5) > 140°^{a}$
$(\angle O6 \cdots H - O2) > 140°^{a}$

Hexamer Ring



$110° < (\angle O1 - O2 - O3) < 130°$
$110° < (\angle O2 - O3 - O4) < 130°$
$110° < (\angle O3 - O4 - O5) < 130°$
$110° < (\angle O4 - O5 - O6) < 130°$
$110° < (\angle O5 - O6 - O1) < 130°$
$|\angle O1 - O2 - O3 - O4| \leq 30°$
$|\angle O1 - O6 - O5 - O4| \leq 30°$
$|\angle O2 - O1 - O4 - O5| \geq 150°$
$|\angle O6 - O1 - O4 - O3| \geq 150°$

(a) This angle is used if H-bond angle, θ, is larger than this angle. Otherwise, θ is used.

6.0 EXAMPLES

In this section, sample input and output files will be discussed for the bulk water, water/Formamide binary mixture, Water/Uranyl mixture, and Water/Methanol/Ethanol ternary mixture.

6.1 BULK WATER

A single snapshot from the molecular dynamics simulation of bulk water at room temperature will be used for the purpose of illustrating the *ChemNetworks* input file preparation. A detailed explanation of the input and output files will be provided.

Input File (Water.input)

The blue numbers on the left of this sample input file are line numbers for reference in this manual, and are not used in practice.

```
1  [NUMBER OF SOLVENT TYPES] 1
2  [NUMBER OF SOLUTE TYPES] 0
3  [NUMBER OF ATOMS IN SOLVENT1] 3
4  O 1
5  H 2
6  H 3
7  [PERIODIC BOUNDARY CONDITIONS] 1
8  [BOX XSIDE] 18.643
9  [BOX YSIDE] 18.643
10 [BOX ZSIDE] 18.643
11 [GRAPH SOLVENT1 SOLVENT1] 1
12 [SOLVENT1 SOLVENT1 HBOND DISTANCE] 4
13 1 2 2.50
14 1 3 2.50
15 2 1 2.50
16 3 1 2.50
17 [SOLVENT1 SOLVENT1 HBOND ANGLE] 4
18 1 2
19 1 2 1 150.0
20 1 2
21 1 3 1 150.0
22 2 1
23 1 2 1 150.0
24 2 1
25 1 3 1 150.0
26 [GRAPH SOLVENT2 SOLVENT2] 0
```

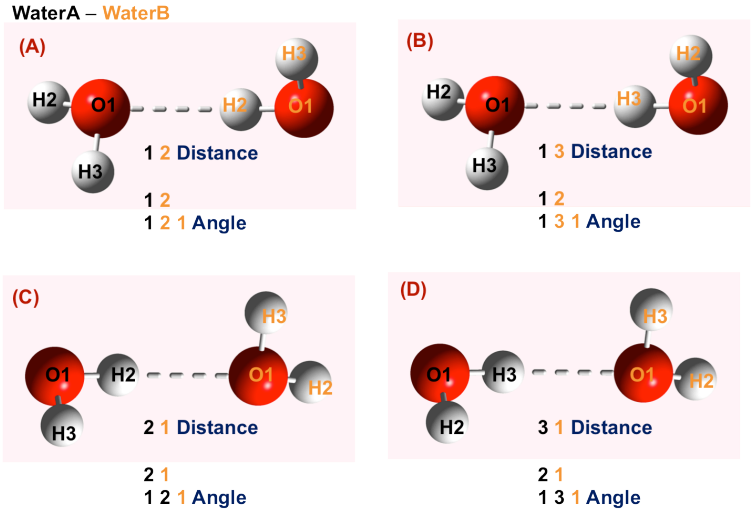


Figure 1. Water-Water hydrogen bonding interactions through all possible O...H pairs.

```
27 [GRAPH SOLVENT3 SOLVENT3] 0
28 [GRAPH SOLVENT1 SOLVENT2] 0
29 [GRAPH SOLVENT1 SOLVENT3] 0
30 [GRAPH SOLVENT1 SOLUTE1] 0
31 [GRAPH SOLVENT1 SOLUTE2] 0
32 [GRAPH SOLVENT2 SOLVENT3] 0
33 [GRAPH SOLVENT2 SOLUTE1] 0
34 [GRAPH SOLVENT2 SOLUTE2] 0
35 [GRAPH SOLVENT3 SOLUTE1] 0
36 [GRAPH SOLVENT3 SOLUTE2] 0
37 [GRAPH SOLVENT1 SOLVENT2 SOLVENT3] 0
38 [PRINT NUMBER OF NODES] 1
39 [GEODESICS GD] 1
40 [GD SOLVENT1] 1
41 [GD SOLVENT1 EUCLIDEAN DISTANCE] 1
42 [GD SOLVENT1 EUCLIDEAN REFERENCE] 1
43 [GD SOLVENT2] 0
44 [GD SOLVENT2 EUCLIDEAN DISTANCE] 0
45 [GD SOLVENT2 EUCLIDEAN REFERENCE] 0
46 [GD SOLVENT3] 0
47 [GD SOLVENT3 EUCLIDEAN DISTANCE] 0
48 [GD SOLVENT3 EUCLIDEAN REFERENCE] 0
49 [GD SOLVENT1 SOLVENT2] 0
50 [GD SOLVENT1 SOLVENT2 EUCLIDEAN DISTANCE] 0
51 [GD SOLVENT1 SOLVENT2 EUCLIDEAN REFERENCE] 0 0
52 [GD SOLVENT1 SOLVENT3] 0
53 [GD SOLVENT1 SOLVENT3 EUCLIDEAN DISTANCE] 0
54 [GD SOLVENT1 SOLVENT3 EUCLIDEAN REFERENCE] 0 0
55 [GD SOLVENT2 SOLVENT3] 0
56 [GD SOLVENT2 SOLVENT3 EUCLIDEAN DISTANCE] 0
57 [GD SOLVENT2 SOLVENT3 EUCLIDEAN REFERENCE] 0 0
58 [GD SOLVENT1 SOLVENT2 SOLVENT3] 0
59 [GD SOLVENT1 SOLVENT2 SOLVENT3 EUCLIDEAN DISTANCE] 0
60 [GD SOLVENT1 SOLVENT2 SOLVENT3 EUCLIDEAN REFERENCE] 0 0 0
61 [SOLUTE1 WATER DIPOLE ORIENTATIONS] 0
62 [SOLUTE2 WATER DIPOLE ORIENTATIONS] 0
63 [SOLVENT WATER DIPOLE ORIENTATIONS] 0 0 0
64 [WATER STRUCTURES] 1 1
65 [WATER HEXAMER RING] 0
66 [WATER HEXAMER BOOK] 0
67 [WATER HEXAMER PRISM] 0
68 [WATER HEXAMER CAGE] 0
69 [WATER HEXAMER BAG] 0
70 [WATER HEXAMER BOAT] 0
71 [WATER HEXAMER CHAIR] 1
72 [WATER HEXAMER PRISMBOOK] 0
```

```
73 [WATER PENTAMER SEARCH] 0
74 [WATER TETRAMER SEARCH] 1
75 [WATER TRIMER SEARCH] 0
76 [WATER ISOLATED STRUCTURES] 0
```

Line 1: Required keyword for the number of major species of interest. Integer value is 1, here, as we will construct and analyze the H-bond network of bulk water only.

Line 2: Required keyword for the number of minor species of interest. Integer value is 0, here, as bulk water is the only species of interest.

Line 3: Number of atoms in a water molecule is 3.

Lines 4 to 6: Atom labels used in the .xyz file for the water, and their position/order numbers next to labels. Here, the water molecule is described as OHH; oxygen is the first, with label "O" and then the hydrogens with the label "H" for both. **There are absolutely no restrictions on the atom labels used and the way they are ordered to describe the molecule of interest.**

Line 7: Periodic boundary conditions are requested to be taken into account across the boundaries of the rectangular box of which has dimensions must be provided.

Lines 8 to 10: X-, Y-, and Z-dimensions of the box, respectively. Must be given in angstroms. Here, we have a cubic box of side-length 18.643Å.

Line 11: The water H-bond graph/network is requested to be constructed.

Line 12: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters.

Lines 13 to 16: Distance criteria for the edge formation. The cutoff distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is set to 2.5 Å (line 13). The cutoff distance between the 1st atom of water A (O-atom) and the 3rd atom of water B (H-atom) is set to 2.5 Å (line 14). The cutoff distance between the 2nd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.5 Å (line 15). The cutoff distance between the 3rd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.5 Å (line 16).

Line 17: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters. The value here must match the one entered in line 12.

Lines 18 to 25: Angle criteria for the edge formation. Angle requires 3 atoms, 2 of which from one water and 1 atom from the other water. The H-atom is always at the middle (OHO) in the angle definitions. When 1 atom is from the 1st molecule and 2 atoms from the 2nd molecule "1 2" combination is used (lines 18 and 20), whereas when 2 atoms are from the 1st molecule and 1 atom from the 2nd molecule "2 1" combination is used (lines 22 and 24). The order of angle criteria must be consistent with the order of distance criteria. Line 18: 1 atom from water A, and 2 atoms from water B. Line 19: The angle cutoff between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), these two atoms were used in line 13, and the 1st atom of water B (O-atom) is set to 150.0°. Therefore, whenever the distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is **less than** 2.5 Å, and the angle between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), and the 1st atom of water B (O-atom) is **larger than** 150.0°, an edge will be formed between the water A and water B. As you can see, lines 13 and 19 define geometric criteria (distance and angle) together. Similarly, lines 14 and 21; lines 15 and 23; and lines 16 and 25 match in defining the geometric criteria. All pairs of water molecules satisfying one of these criteria will be considered H-bonded and printed as pairs in the output file ending with *.Graph*.

Lines 26 to 37: All other keywords (starting with GRAPH) for graph/network construction are set to 0, as bulk water is the only chemical system at hand.

Line 38: The number of vertices/nodes (water molecules in this example) making up the graph/network will be printed in an output file ending with *.GraphNumnodes*.

Line 39: The geodesic distance (gd) matrix and Euclidean distance for the water H-bond network will be calculated.

Line 40: The gd matrix for the network obtained from the keyword in line 11 will be calculated.

Line 41: Euclidean distance between the terminal water molecules of all geodesic pathways within the water H-bond network will be calculated.

Line 42: Euclidean distance between the 1st atom (O-atom) of terminal water molecules of all geodesic pathways will be calculated.

Lines 43 to 60: All other keywords (starting with GD) for gd matrix construction for other graph/network types are set to 0, as bulk water is the only chemical system at hand.

Lines 61 to 62: Water dipole moment orientations with respect to a solute are requested using these keywords. In this example, there are no solutes present.

Line 63: Water dipole moment orientations with respect to other solvent molecules are requested using this keyword. In this example, there are no other solvents.

Line 64: The search for water oligomers will be done. Water is the 1st "solvent" defined above.

Lines 65 to 75: The search only for the chair hexamers, and cyclic water tetramers will be done.

Line 76: The isolated oligomers (that are isolated from the rest of the network) won't be searched.

Running: ./ChemNetworks.exe Water.input water.xyz

Output Files

Water.input.water.xyz.Water.Cyclic.Tetramers
Water.input.water.xyz.Water.Hexamer.Chairs
Water.input.water.xyz.water.xyz.Graph
Water.input.water.xyz.water.xyz.GraphGeod
Water.input.water.xyz.water.xyz.GraphNumnodes
water.xyz.geocard
water.xyz.geopath

The list of all output file names is shown above. Let's take a closer look at these output files.

Water H-bond network (edge-list) [GRAPH SOLVENT1 SOLVENT1]

Water.input.water.xyz.water.xyz.Graph

```
  1
 57
  1
132
  2
 55
  :
  :
212
139
  1
120
  7
 66
```

Water.input.water.xyz.water.xyz.GraphGeod

```
1 57 0 0 0
1 132 0 0 0
2 55 0 0 0
…
…
212 139 0 1 0
1 120 0 0 1
7 66 0 0 1
…
```

Edge from PBC (+z image)

The two output files above show the list of the pairs of water molecules (edge list) that are

hydrogen bonded based on the user defined geometric criteria. The output file on the left lists the water ID numbers such that the water # 1 and water #57 are H-bonded, etc. The one on the right-hand side shows not only the edge list but also the information about whether the edge is formed through the PBC boundaries. For example, an edge is formed between the water #1 and the water #120 across the PBC boundary (+z image); water #1 is in the central MD box while the image of water #120 is in the +z image of the central box. The output file ending with *.GraphGeod* is used to calculate the geodesic distance matrix, with the periodic boundary conditions are fully taken into account.

The number of vertices/nodes making up the graph/network is printed in the output file, *Water.input.water.xyz.water.xyz.GraphNumnodes,* as 216, which is the total number of water molecules in the box.

The geodesic distance matrix is printed in the output file named *water.xyz.geopath*. The first and the second integers at each line represents the water ID numbers, the third integer is the minimum number of contiguous H-bonds that connects the two waters. The actual geodesic path involving the participating waters is also printed. The output file named *water.xyz.geocard* holds the same information in addition to the Euclidean distance printed at the end of each line. **Note that** the geodesic distance matrix is a symmetric N x N matrix, i.e. the shortest path from vertex *i* to vertex *j* is the same as the shortest path from vertex *j* to vertex *i*. Therefore, to prevent double counting only the matrix elements from the upper triangle are printed in the output files ending with *.geopath* and *.geocard*.

Water H-bond network (geodesic distance matrix) [GD SOLVENT1]

water.xyz.geopath

```
1 2 8 path: 1 57 33 7 151 153 61 103 2
1 3 5 path: 1 57 126 112 134 3
1 4 5 path: 1 57 126 112 46 4
1 5 9 path: 1 132 186 119 136 121 170 179 102 5
1 6 7 path: 1 57 33 200 56 37 27 6
1 7 3 path: 1 57 33 7
1 8 4 path: 1 132 186 119 8
1 9 4 path: 1 120 59 193 9
          .
          .
          .
214 216 8 path: 214 146 129 81 142 84 83 71 216
215 216 6 path: 215 131 140 12 18 45 216
```

The waters that are forming water oligomers, cyclic tetramer and chair hexamer here, are printed in the following output files:

Water H-bond network (water oligomers) [WATER STRUCTURES]

Water.input.water.xyz.Water.Cyclic.Tetramers

7 33 77 70

Water.input.water.xyz.Water.Hexamer.Chairs

3 29 100 41 16 195

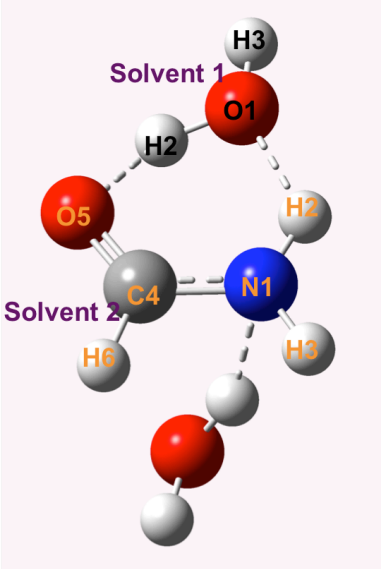
6.2 WATER/FORMAMIDE BINARY MIXTURE

ChemNetworks input and output files will be discussed for the binary mixture of solvents water and formamide (CH_3NO). Mixture of 200 H_2O molecules and 50 CH_3NO molecules are put inside a rectangular box of sides 22 Å x 27 Å x 24 Å. In this example, three types of graphs will be constructed: (i) Water only network, (ii) Formamide only network, and (iii) the entire network including (i), (ii), and the network between the water and formamide.

Input File (WaterFormamide.input)

The blue numbers on the left of this sample input file are line numbers for reference in this manual, and are not used in practice.

```
1  [NUMBER OF SOLVENT TYPES] 2
2  [NUMBER OF SOLUTE TYPES] 0
3  [NUMBER OF ATOMS IN SOLVENT1] 3
4   O 1
5   H 2
6   H 3
7  [NUMBER OF ATOMS IN SOLVENT2] 6
8  N 1
9  H 2
10 H 3
11 C 4
12 O 5
13 H 6
14 [PERIODIC BOUNDARY CONDITIONS] 1
15 [BOX XSIDE] 22.0
16 [BOX YSIDE] 27.0
17 [BOX ZSIDE] 24.0
18 [GRAPH SOLVENT1 SOLVENT1] 1
19 [SOLVENT1 SOLVENT1 HBOND DISTANCE] 4
20 1 2 2.50
21 1 3 2.50
22 2 1 2.50
23 3 1 2.50
24 [SOLVENT1 SOLVENT1 HBOND ANGLE] 4
25 1 2
26 1 2 1 145.0
27 1 2
28 1 3 1 145.0
29 2 1
30 1 2 1 145.0
31 2 1
32 1 3 1 145.0
```



```
33 [GRAPH SOLVENT2 SOLVENT2] 1
34 [SOLVENT2 SOLVENT2 HBOND DISTANCE] 4
35 5 2 2.85
36 5 3 2.85
37 2 5 2.85
38 3 5 2.85
39 [SOLVENT2 SOLVENT2 HBOND ANGLE] 4
40 1 2
41 5 2 1 125
42 1 2
43 5 3 1 125
44 2 1
45 1 2 5 125
46 2 1
47 1 3 5 125
48 [GRAPH SOLVENT3 SOLVENT3] 0
49 [GRAPH SOLVENT1 SOLVENT2] 1
50 [SOLVENT1 SOLVENT2 HBOND DISTANCE] 6
51 1 2 2.75
52 1 3 2.75
53 2 1 2.75
54 3 1 2.75
55 2 5 2.75
56 3 5 2.75
57 [SOLVENT1 SOLVENT2 HBOND ANGLE] 6
58 1 2
59 1 2 1 125
60 1 2
61 1 3 1 125
62 2 1
63 1 2 1 125
64 2 1
65 1 3 1 125
66 2 1
67 1 2 5 125
68 2 1
69 1 3 5 125
70 [GRAPH SOLVENT1 SOLVENT3] 0
71 [GRAPH SOLVENT1 SOLUTE1] 0
72 [GRAPH SOLVENT1 SOLUTE2] 0
73 [GRAPH SOLVENT2 SOLVENT3] 0
74 [GRAPH SOLVENT2 SOLUTE1] 0
75 [GRAPH SOLVENT2 SOLUTE2] 0
76 [GRAPH SOLVENT3 SOLUTE1] 0
77 [GRAPH SOLVENT3 SOLUTE2] 0
78 [GRAPH SOLVENT1 SOLVENT2 SOLVENT3] 0
```

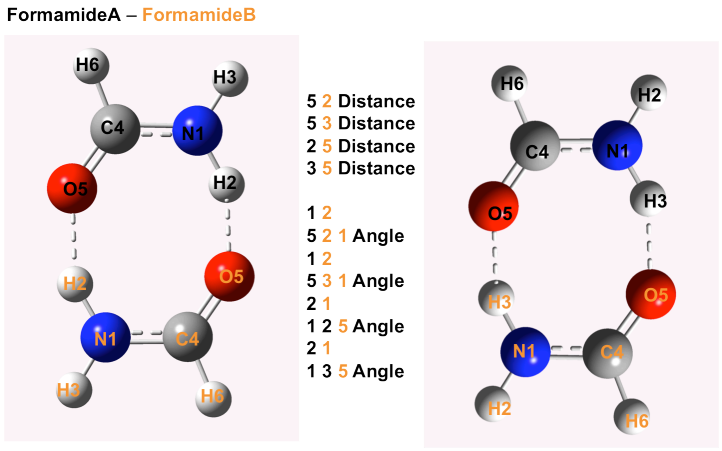


Figure 2. Formamide-Formamide hydrogen bonding interactions that are used in edge formation.

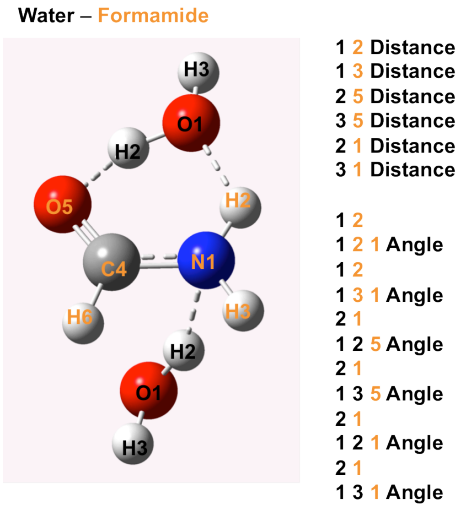


Figure 3. Water-Formamide hydrogen bonding interactions that are used in edge formation.

```
79 [PRINT NUMBER OF NODES] 1
80 [GEODESICS GD] 0
81 [SOLUTE1 WATER DIPOLE ORIENTATIONS] 0
82 [SOLUTE2 WATER DIPOLE ORIENTATIONS] 0
83 [SOLVENT WATER DIPOLE ORIENTATIONS] 0 0 0
84 [WATER STRUCTURES] 0
```

Line 1: Required keyword for the number of major species of interest. Integer value is 2, here, as we have two chemical species.

Line 2: No solute species are present. Therefore, value is set to 0.

Line 3: Number of atoms in a water molecule ("solvent 1") is 3.

Lines 4 to 6: Atom labels used in the .xyz file for the water, and their position/order numbers next to labels. Here, the water molecule is described as OHH; oxygen is the first, with label "O" and then the hydrogens with the label "H" for both.

Line 7: Number of atoms in a formamide molecule ("solvent 2") is 6.

Lines 8 to 13: Atom labels used in the .xyz file for the formamide, and their position/order numbers next to labels. Here, the formamide molecule is described as NHHCOH. **There are absolutely no restrictions on the atom labels used and the way they are ordered to describe the molecule of interest.**

Line 14: Periodic boundary conditions are requested to be taken into account across the boundaries of the rectangular box of which has dimensions must be provided.

Lines 15 to 17: X-, Y-, and Z-dimensions of the box, respectively. In this example, we have a rectangular box of side-lengths 22 Å x 27 Å x 24 Å.

Line 18: The water H-bond graph/network is requested to be constructed.

Line 19: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters.

Lines 20 to 23: Distance criteria for the edge formation. The cutoff distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is set to 2.5 Å (line 20). The cutoff distance between the 1st atom of water A (O-atom) and the 3rd atom of water B (H-atom) is set to

2.5 Å (line 21). The cutoff distance between the 2nd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.5 Å (line 22). The cutoff distance between the 3rd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.5 Å (line 23).

Line 24: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters. The value here must match the one entered in line 19.

Lines 25 to 32: Angle criteria for the edge formation. Angle requires 3 atoms, 2 of which from one water and 1 atom from the other water. The H-atom is always at the middle (OHO) in the angle definitions. When 1 atom is from the 1st molecule and 2 atoms from the 2nd molecule "1 2" combination is used (lines 25 and 27), whereas when 2 atoms are from the 1st molecule and 1 atom from the 2nd molecule "2 1" combination is used (lines 29 and 31). The order of angle criteria must be consistent with the order of distance criteria. Line 25: 1 atom from water A, and 2 atoms from water B. Line 26: The angle cutoff between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), these two atoms were used in line 20, and the 1st atom of water B (O-atom) is set to 145.0°. Therefore, whenever the distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is **less than** 2.5 Å, and the angle between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), and the 1st atom of water B (O-atom) is **larger than** 145.0°, an edge will be formed between the water A and water B. As you can see, lines 20 and 26 define geometric criteria (distance and angle) together. Similarly, lines 21 and 28; lines 22 and 30; and lines 23 and 32 match in defining the geometric criteria. All pairs of water molecules satisfying one of these criteria will be considered H-bonded and printed as pairs in the output file ending with *.Graph*.

Line 33: Formamide only H-bond graph/network is requested to be constructed.

Line 34: The number of H-bond interactions to be used to describe the edges (H-bond) between two formamide molecules (see Figure 2). In this example, the hydrogen bond interactions involving only the O...H pairs are taken into account.

Lines 35 to 38: Distance criteria for the edge formation. The cutoff distance between the 5th atom of formamide A (O-atom) and the 2nd atom of formamide B (H-atom) is set to 2.85 Å (line 35). The cutoff distance between the 5th atom of formamide A (O-atom) and the 3rd atom of formamide B (H-atom) is set to 2.85 Å (line 36). The cutoff distance between the 2nd atom of formamide A (H-atom) and the 5th atom of formamide B (O-atom) is set to 2.85 Å (line 37). The cutoff distance between the 3rd atom of formamide A (H-atom) and the 5th atom of formamide B (O-atom) is set to 2.85 Å (line 38).

Line 39: The number of H-bond interactions to be used to describe the edges between the two formamide molecules (see Figure 2). The value here must match the one entered in line 34.

Lines 40 to 47: Angle criteria for the edge formation. Angle requires 3 atoms, 2 of which from one formamide and 1 atom from the other formamide. The H-atom is always at the middle in the angle definitions. When 1 atom is from the 1st molecule and 2 atoms from the 2nd molecule "1 2" combination is used (lines 40 and 42), whereas when 2 atoms are from the 1st molecule and 1 atom from the 2nd molecule "2 1" combination is used (lines 44 and 46). The order of angle criteria must be consistent with the order of distance criteria. Line 40: 1 atom from formamide A, and 2 atoms from formamide B. Line 41: The angle cutoff between the 5th atom of formamide A (O-atom) and the 2nd atom of formamide B (H-atom), these two atoms were used in line 35, and the 1st atom of formamide B (N-atom) is set to 125.0°. Therefore, whenever the distance between the 5th atom of formamide A (O-atom) and the 2nd atom of formamide B (H-atom) is **less than** 2.5 Å, and the angle between the 5th atom of formamide A (O-atom) and the 2nd atom of formamide B (H-atom), and the 1st atom of formamide B (N-atom) is **larger than** 125.0°, an edge will be formed between the formamide A and formamide B. As you can see, lines 35 and 41 define geometric criteria (distance and angle) together. Similarly, lines 36 and 43; lines 37 and 45; and lines 38 and 47 match in defining the geometric criteria. All pairs of formamide molecules satisfying one of these criteria will be considered H-bonded and printed as pairs in the output file ending with *.Graph*.

Line 48: Value is set to 0, as there is no third type species present in the chemical system.

Line 49: The entire network formed between the water and formamide including water-water, formamide-water, and water-formamide interactions/edges will be constructed. If the network of only the water-formamide interactions/edges is needed, then only the keyword in line 49 should be activated and the keywords in lines 18 and 33 must be turned off.

Line 50: The number of H-bond interactions to be used to describe the edges between water-formamide molecules (see Figure 3).

Lines 51 to 56: Distance criteria for the edge formation. The cutoff distance between the 1st atom of water (O-atom) and the 2nd atom of formamide (H-atom) is set to 2.75 Å (line 51). The cutoff distance between the 1st atom of water (O-atom) and the 3rd atom of formamide (H-atom) is set to 2.75 Å (line 52). The cutoff distance between the 2nd atom of water (H-atom) and the 5th atom of formamide (O-atom) is set to 2.75 Å (line 55). The cutoff distance between the 3rd atom of water (H-atom) and the 5th atom of formamide (O-atom) is set to 2.75 Å (line 56). The cutoff distance between the 2nd atom of water (H-atom) and the 1st atom of formamide (N-atom) is set

to 2.75 Å (line 53). The cutoff distance between the 3rd atom of water (H-atom) and the 1st atom of formamide (N-atom) is set to 2.75 Å (line 54).

Line 57: The number of H-bond interactions to be used to describe the edges between the water-formamide molecules (see Figure 3). The value here must match the one entered in line 50.

Lines 58 to 69: Angle criteria for the edge formation. Angle requires 3 atoms, 2 of which from one molecule and 1 atom from the other molecule. The H-atom is always at the middle in the angle definitions. When 1 atom is from the 1st molecule and 2 atoms from the 2nd molecule "1 2" combination is used (lines 58 and 60), whereas when 2 atoms are from the 1st molecule and 1 atom from the 2nd molecule "2 1" combination is used (lines 62, 64, 66, and 68). The order of angle criteria must be consistent with the order of distance criteria. Line 58: 1 atom from water, and 2 atoms from formamide. Line 59: The angle cutoff between the 1st atom of water (O-atom) and the 2nd atom of formamide (H-atom), these two atoms were used in line 51, and the 1st atom of formamide (N-atom) is set to 125.0°. Therefore, whenever the distance between the 1st atom of water (O-atom) and the 2nd atom of formamide (H-atom) is **less than** 2.75 Å, and the angle between the 1st atom of water (O-atom) and the 2nd atom of formamide (H-atom), and the 1st atom of formamide (N-atom) is **larger than** 125.0°, an edge will be formed between the water and formamide molecules. As you can see, lines 51 and 59 define geometric criteria (distance and angle) together. Similarly, lines 52 and 61; lines 53 and 63; lines 54 and 65; lines 55 and 67; and lines 56 and 69 match in defining the geometric criteria. All pairs of water-formamide molecules satisfying one of these criteria will be considered H-bonded and printed as pairs in the output file ending with *.Graph*.

Lines 70 to 78: All other keywords (starting with GRAPH) for graph/network construction are set to 0.

Line 79: The number of vertices/nodes (water molecules in this example) making up all requested graphs will be printed in an output files ending with *.GraphNumnodes*.

Line 80: No geodesic distance analysis is requested.

Lines 81 to 83: No water dipole orientation calculations are requested.

Line 84: No water oligomer search is requested.

Running: ./ChemNetworks.exe WaterFormamide.input water.xyz formamide.xyz

Output Files

The list of all output file names is shown below.

Water only [GRAPH SOLVENT1 SOLVENT1]

WaterFormamide.input.water.xyz.water.xyz.Graph
WaterFormamide.input.water.xyz.water.xyz.GraphGeod
WaterFormamide.input.water.xyz.water.xyz.GraphNumnodes

Formamide only [GRAPH SOLVENT2 SOLVENT2]

WaterFormamide.input.formamide.xyz.formamide.xyz.Graph
WaterFormamide.input.formamide.xyz.formamide.xyz.GraphGeod
WaterFormamide.input.formamide.xyz.formamide.xyz.GraphNumnodes

Water-Formamide edges [GRAPH SOLVENT1 SOLVENT2]

WaterFormamide.input.water.xyz.formamide.xyz.Graph
WaterFormamide.input.water.xyz.formamide.xyz.GraphGeod
WaterFormamide.input.water.xyz.formamide.xyz.GraphNumnodes

[GRAPH SOLVENT1 SOLVENT1]

WaterFormamide.input.water.xyz.water.xyz.GraphGeod

3 28 0 0 0
7 127 0 0 0
12 93 0 0 0
19 84 0 0 0
…

Water- Water

[GRAPH SOLVENT2 SOLVENT2]

WaterFormamide.input.formamide.xyz.formamide.xyz.GraphGeod

7 25 0 0 0
9 48 0 0 0
16 42 0 0 0
…

Formamide - Formamide

Output files above list the edges for the water only and formamide only network. The edge-list for the entire network is printed in *WaterFormamide.input.water.xyz.formamide.xyz.Graph*. The same list is also shown in *WaterFormamide.input.water.xyz.formamide.xyz.GraphGeod* file:

Water-Formamide (entire network) [GRAPH SOLVENT1 SOLVENT2]

WaterFormamide.input.water.xyz.formamide.xyz.GraphGeod

```
3 28 0 0 0
7 127 0 0 0
12 93 0 0 0
19 84 0 0 0
…
207 225 0 0 0
209 248 0 0 0
216 242 0 0 0
5 221 0 0 0
10 210 0 0 0
20 239 0 0 0
…
31 231 0 1 0
194 249 0 1 0
```

Water- Water

Formamide - Formamide

Water - Formamide

Edge from PBC (+y image)

WaterFormamide.input.water.xyz.formamide.xyz.GraphNumnodes

250

The numbering of vertices/nodes in the water only network runs from 1 to 200, which is the total number of water molecules. The numbering of vertices/nodes in the formamide only network runs from 1 to 50, which is the total number of formamide molecules. However, the numbering runs from 1 to 250 in the entire network, such that waters from 1 to 200 and formamides from 201 to 250. Note that the first “solvent” is water, and the second is formamide in the input file.

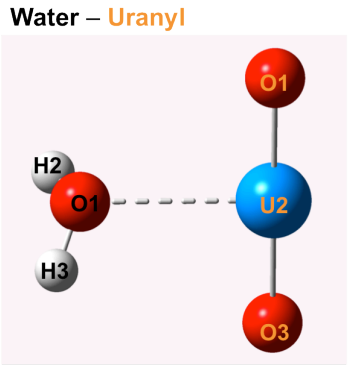
6.3 WATER/URANYL MIXTURE

ChemNetworks input and output files will be discussed for the mixture of solvent water and solute uranyl, $(UO_2)^{2+}$. Mixture of 215 H_2O molecules and 3 uranyl molecules are put inside a rectangular box of sides 22.0 Å x 23.2 Å x 21.9 Å. In this example, two types of graphs will be constructed: (i) Water only network, and (ii) the entire network including (i), and the network between the water and uranyl.

Input File (WaterUranyl.input)

The blue numbers on the left of this sample input file are line numbers for reference in this manual, and are not used in practice.

```
1  [NUMBER OF SOLVENT TYPES] 1
2  [NUMBER OF SOLUTE TYPES] 1
3  [NUMBER OF ATOMS IN SOLVENT1] 3
4   O 1
5   H 2
6   H 3
7  [NUMBER OF ATOMS IN SOLUTE1] 3
8  O 1
9  U 2
10 O 3
11 [PERIODIC BOUNDARY CONDITIONS] 1
12 [BOX XSIDE] 22.0
13 [BOX YSIDE] 23.2
14 [BOX ZSIDE] 21.9
15 [GRAPH SOLVENT1 SOLVENT1] 1
16 [SOLVENT1 SOLVENT1 HBOND DISTANCE] 4
17 1 2 2.75
18 1 3 2.75
19 2 1 2.75
20 3 1 2.75
21 [SOLVENT1 SOLVENT1 HBOND ANGLE] 4
22 1 2
23 1 2 1 145
24 1 2
25 1 3 1 145
26 2 1
27 1 2 1 145
28 2 1
29 1 3 1 145
30 [GRAPH SOLVENT2 SOLVENT2] 0
31 [GRAPH SOLVENT3 SOLVENT3] 0
32 [GRAPH SOLVENT1 SOLVENT2] 0
```



```
33 [GRAPH SOLVENT1 SOLVENT3] 0
34 [GRAPH SOLVENT1 SOLUTE1] 1
35 [SOLVENT1 SOLUTE1 CUTOFF] 1
36 1 2 2.70
37 [GRAPH SOLVENT1 SOLUTE2] 0
38 [GRAPH SOLVENT2 SOLVENT3] 0
39 [GRAPH SOLVENT2 SOLUTE1] 0
40 [GRAPH SOLVENT2 SOLUTE2] 0
41 [GRAPH SOLVENT3 SOLUTE1] 0
42 [GRAPH SOLVENT3 SOLUTE2] 0
43 [GRAPH SOLVENT1 SOLVENT2 SOLVENT3] 0
44 [PRINT NUMBER OF NODES] 1
45 [GEODESICS GD] 0
46 [SOLUTE1 WATER DIPOLE ORIENTATIONS] 1 1
47 [SOLUTE2 WATER DIPOLE ORIENTATIONS] 0
48 [SOLVENT WATER DIPOLE ORIENTATIONS] 0 0 0
49 [WATER STRUCTURES] 0
```

Line 1: Required keyword for the number of major species of interest. Integer value is 1, here, water is the only "solvent".

Line 2: The Uranyl is the "solute" in this example. Integer value is 1.

Line 3: Number of atoms in a water molecule ("solvent 1") is 3.

Lines 4 to 6: Atom labels used in the .xyz file for the water, and their position/order numbers next to labels. Here, the water molecule is described as OHH; oxygen is the first, with label "O" and then the hydrogens with the label "H" for both.

Line 7: Number of atoms in a uranyl molecule ("solute 1") is 3.

Lines 8 to 10: Atom labels used in the .xyz file for the uranyl, and their position/order numbers next to labels. Here, the uranyl molecule is described as OUO; uranium is the second, with label "U" and the two oxygens with the label "O" for both at the first and the third positions. **There are absolutely no restrictions on the atom labels used and the way they are ordered to describe the molecule of interest.**

Line 11: Periodic boundary conditions will be taken into account across the boundaries of the rectangular box of which has dimensions must be provided.

Lines 12 to 14: X-, Y-, and Z-dimensions of the box, respectively. In this example, we have a rectangular box of side-lengths 22.0 Å x 23.2 Å x 21.9 Å.

Line 15: The water H-bond graph/network will be constructed.

Line 16: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters.

Lines 17 to 20: Distance criteria for the edge formation. The cutoff distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is set to 2.75 Å (line 17). The cutoff distance between the 1st atom of water A (O-atom) and the 3rd atom of water B (H-atom) is set to 2.75 Å (line 18). The cutoff distance between the 2nd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.75 Å (line 19). The cutoff distance between the 3rd atom of water A (H-atom) and the 1st atom of water B (O-atom) is set to 2.75 Å (line 20).

Line 21: The number of H-bond interactions to be used to describe the edges (H-bond) between two water molecules (see Figure 1). Here, there are four possible O...H pairs for the two waters. The value here must match the one entered in line 16.

Lines 22 to 29: Angle criteria for the edge formation. Angle requires 3 atoms, 2 of which from one water and 1 atom from the other water. The H-atom is always at the middle (OHO) in the angle definitions. When 1 atom is from the 1st molecule and 2 atoms from the 2nd molecule "1 2" combination is used (lines 22 and 24), whereas when 2 atoms are from the 1st molecule and 1 atom from the 2nd molecule "2 1" combination is used (lines 26 and 28). The order of angle criteria must be consistent with the order of distance criteria. Line 22: 1 atom from water A, and 2 atoms from water B. Line 23: The angle cutoff between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), these two atoms were used in line 17, and the 1st atom of water B (O-atom) is set to 145.0°. Therefore, whenever the distance between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom) is **less than** 2.75 Å, and the angle between the 1st atom of water A (O-atom) and the 2nd atom of water B (H-atom), and the 1st atom of water B (O-atom) is **larger than** 145.0°, an edge will be formed between the water A and water B. As you can see, lines 17 and 23 define geometric criteria (distance and angle) together. Similarly, lines 18 and 25; lines 19 and 27; and lines 20 and 29 match in defining the geometric criteria. All pairs of water molecules satisfying one of these criteria will be considered H-bonded and printed as pairs in the output file ending with *.Graph*.

Lines 30 to 33: These keywords starting with GRAPH are set to 0.

Line 34: The entire graph including water H-bond network, and the network between the water and uranyl will be constructed. For a network containing only the water-uranyl interactions/edges, the keyword in line 15 must be turned off.

Line 35: The number of criteria to be set to define edge between water and uranyl molecules.

Line 36: The cutoff distance between the 1st atom of water (O-atom) and the 2nd atom of uranyl (U-atom) is set to 2.70 Å. Then, whenever the cutoff is **less than** 2.70 Å *ChemNetworks* will put an edge between the corresponding water and uranyl molecules and the water-uranyl pairs will be printed in an output file ending with .*Graph*.

Lines 37 to 43: These keywords starting with GRAPH are set to 0.

Line 44: The number of vertices/nodes making up the two requested graphs will be printed in the corresponding output files ending with .*GraphNumnodes*.

Line 45: No geodesic distance analysis is requested.

Line 46: Water dipole moment orientations about the Uranyl ions will be determined. Water is the solvent #1. The angle between the dipole moment vector, **P**, of a H_2O molecule and the position vector, **r**, which points from the solute molecule to it (as specified in [SOLVENT1 SOLUTE1 CUTOFF] keyword) and lies on the hydrogen bond between the two, is defined as the dipole angle.

Line 47 to 48: No water dipole orientations about another solute/solvent are requested, as the uranyl ion is the only other type of species in this example.

Line 49: No water oligomer search is requested.

Running: ./ChemNetworks.exe WaterUranyl.input water.xyz uranyl.xyz

Output Files

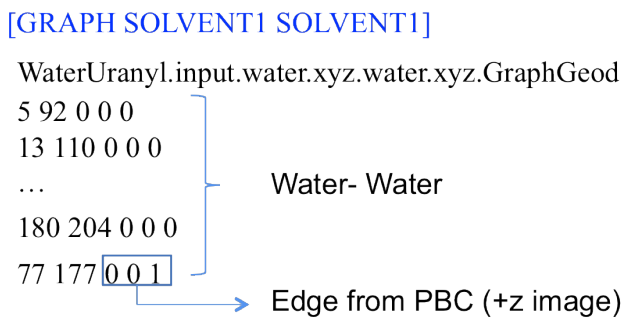
The list of all output file names is shown below.

Water only [GRAPH SOLVENT1 SOLVENT1]
WaterUranyl.input.water.xyz.water.xyz.Graph
WaterUranyl.input.water.xyz.water.xyz.GraphGeod
WaterUranyl.input.water.xyz.water.xyz.GraphNumnodes

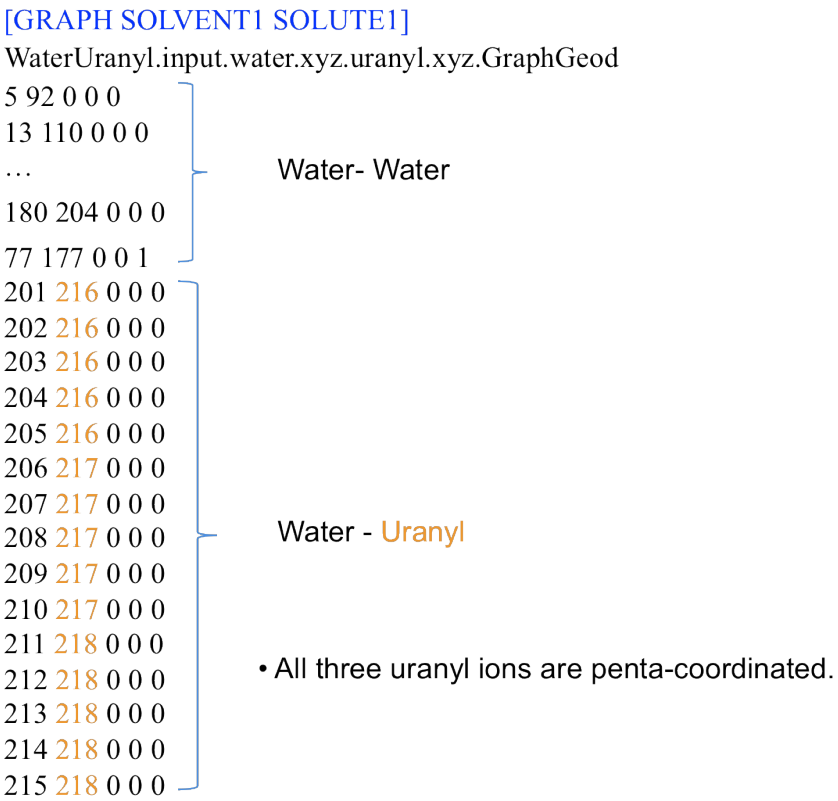
Water-Uranyl egde-list [GRAPH SOLVENT1 SOLUTE1]
WaterUranyl.input.water.xyz.uranyl.xyz.Graph
WaterUranyl.input.water.xyz.uranyl.xyz.GraphGeod
WaterUranyl.input.water.xyz.uranyl.xyz.GraphNumnodes

Water-dipole orientations about Uranyl ions [SOLUTE1 WATER DIPOLE ORIENTATIONS]
WaterUranyl.input.water.xyz.uranyl.xyz.Graph.Dipole.angles

The output file for the edge-list of the water only network is shown below:



The number of vertices/nodes making up the water only graph/network is printed in the output file, *WaterUranyl.input.water.xyz.water.xyz.GraphNumnodes,* as 215, which is the total number of water molecules in this example.



The complete edge-list for the entire network/graph of the water/uranyl mixture is printed in the *WaterUranyl.input.water.xyz.uranyl.xyz.Graph* file. The same edge-list is also shown in the file *WaterUranyl.input.water.xyz.uranyl.xyz.GraphGeod*, shown above.

The numbering of vertices/nodes in the water only network runs from 1 to 215, which is the total number of water molecules, in this example. However, the numbering runs from 1 to 218 in the entire network, such that waters from 1 to 215 and uranyl ions from 216 to 2218. Note that the "solvent" water is defined the first, and the "solute" uranyl ion is defined the next in the input file. The order of species is important when running the *ChemNetworks* and interpreting the output files. In *ChemNetworks*, the solvents precede the solutes.

The orientations of water molecules around the Uranyl ions are determined by the water dipole angles, and printed in the output file ending with the extension .Dipole.angles.

Water dipole orientations [SOLUTE1 WATER DIPOLE ORIENTATIONS]

WaterUranyl.input.water.xyz.uranyl.xyz.Graph.Dipole.angles

Water-Solute: 201 216 Angle: 54.7 Distance: 2.67 WaterAtom-SoluteAtom: 1 2

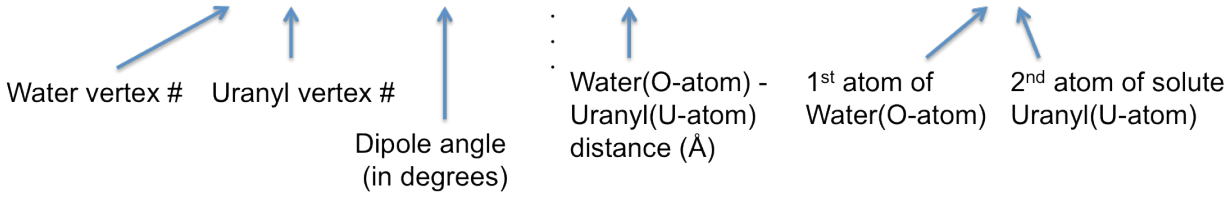
.
.
.

Water-Solute: 206 217 Angle: 54.7 Distance: 2.67 WaterAtom-SoluteAtom: 1 2

.
.
.

Water-Solute: 211 218 Angle: 54.7 Distance: 2.67 WaterAtom-SoluteAtom: 1 2

.
.
.



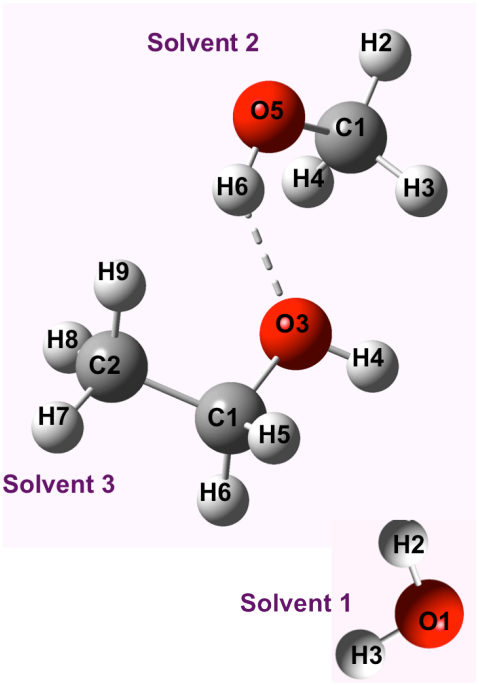
6.4 WATER/METHANOL/ETHANOL TERNARY MIXTURE

Input and output files will be discussed for the ternary mixture of solvents water, methanol (CH_3OH), and ethanol (CH_3CH_2OH). A mixture of 200 water molecules, 70 methanol molecules, and 50 ethanol molecules are put inside a rectangular box. In this example, seven different networks are constructed: water only, methanol only, ethanol only, water-methanol, water-ethanol, methanol-ethanol, and water-methanol-ethanol.

Input File (WaterMethanolEthanol.input)

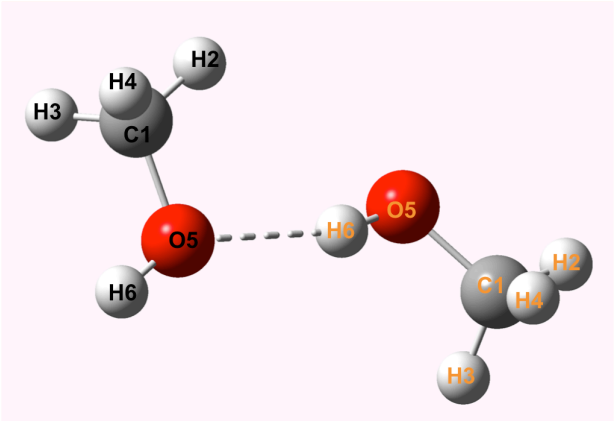
The blue numbers on the left of this sample input file are line numbers for reference in this manual, and are not used in practice.

```
1  [NUMBER OF SOLVENT TYPES] 3
2  [NUMBER OF SOLUTE TYPES] 0
3  [NUMBER OF ATOMS IN SOLVENT1] 3
4   O 1
5   H 2
6   H 3
7  [NUMBER OF ATOMS IN SOLVENT2] 6
8  C 1
9  H 2
10 H 3
11 H 4
12 O 5
13 H 6
14 [NUMBER OF ATOMS IN SOLVENT3] 9
15 C 1
16 C 2
17 O 3
18 H 4
19 H 5
20 H 6
21 H 7
22 H 8
23 H 9
24 [PERIODIC BOUNDARY CONDITIONS] 1
25 [BOX XSIDE] 26.3
26 [BOX YSIDE] 26.2
27 [BOX ZSIDE] 26.1
28 [GRAPH SOLVENT1 SOLVENT1] 1
29 [SOLVENT1 SOLVENT1 HBOND DISTANCE] 4
30 1 2 2.5
31 1 3 2.5
32 2 1 2.5
```

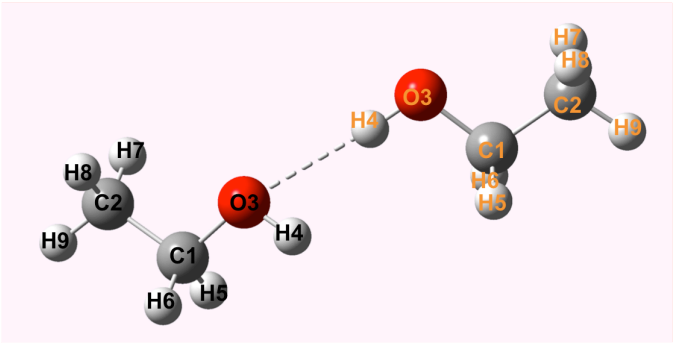


```
3 1 2.5
[SOLVENT1 SOLVENT1 HBOND ANGLE] 4
1 2
1 2 1 145
1 2
1 3 1 145
2 1
1 2 1 145
2 1
1 3 1 145
[GRAPH SOLVENT2 SOLVENT2] 1
[SOLVENT2 SOLVENT2 HBOND DISTANCE] 2
5 6 2.75
6 5 2.75
[SOLVENT2 SOLVENT2 HBOND ANGLE] 2
1 2
5 6 5 125
2 1
5 6 5 125
[GRAPH SOLVENT3 SOLVENT3] 1
[SOLVENT3 SOLVENT3 HBOND DISTANCE] 2
3 4 2.95
4 3 2.95
[SOLVENT3 SOLVENT3 HBOND ANGLE] 2
1 2
3 4 3 125
2 1
3 4 3 125
[GRAPH SOLVENT1 SOLVENT2] 1
[SOLVENT1 SOLVENT2 HBOND DISTANCE] 3
1 6 2.75
2 5 2.75
3 5 2.75
[SOLVENT1 SOLVENT2 HBOND ANGLE] 3
1 2
1 6 5 125
2 1
1 2 5 125
2 1
1 3 5 125
[GRAPH SOLVENT1 SOLVENT3] 1
[SOLVENT1 SOLVENT3 HBOND DISTANCE] 3
1 4 2.95
2 3 2.95
3 3 2.95
[SOLVENT1 SOLVENT3 HBOND ANGLE] 3
```

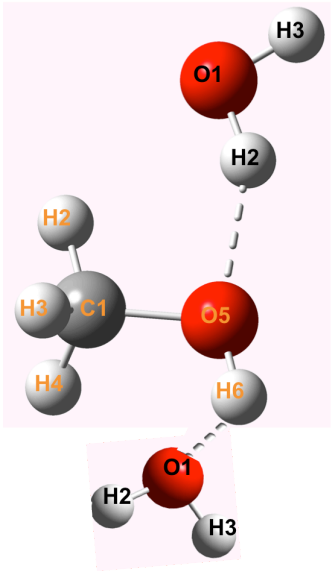
Methanol – **Methanol**



Ethanol – **Ethanol**

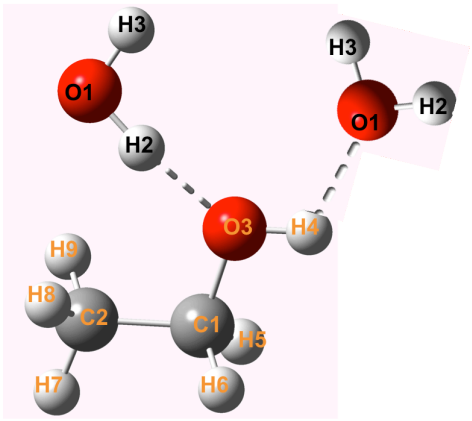


Water – **Methanol**

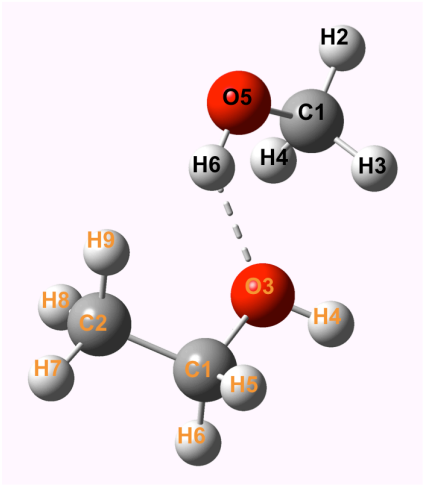


```
1 2
1 4 3 125
2 1
1 2 3 125
2 1
1 3 3 125
[GRAPH SOLVENT1 SOLUTE1] 0
[GRAPH SOLVENT1 SOLUTE2] 0
[GRAPH SOLVENT2 SOLVENT3] 1
[SOLVENT2 SOLVENT3 HBOND DISTANCE] 2
5 4 2.95
6 3 2.95
[SOLVENT2 SOLVENT3 HBOND ANGLE] 2
1 2
5 4 3 125
2 1
5 6 3 125
[GRAPH SOLVENT2 SOLUTE1] 0
[GRAPH SOLVENT2 SOLUTE2] 0
[GRAPH SOLVENT3 SOLUTE1] 0
[GRAPH SOLVENT3 SOLUTE2] 0
[GRAPH SOLVENT1 SOLVENT2 SOLVENT3] 1
[PRINT NUMBER OF NODES] 1
[GEODESICS GD] 0
[SOLUTE1 WATER DIPOLE ORIENTATIONS] 0
[SOLUTE2 WATER DIPOLE ORIENTATIONS] 0
[SOLVENT WATER DIPOLE ORIENTATIONS] 0 0 0
[WATER STRUCTURES] 0
```

Water – **Ethanol**



Methanol – **Ethanol**



General principles used in the input files that are discussed in detail in the preceding sections apply also in this example. Therefore, redundant details will be omitted. Here, several networks will be constructed:

- water only network (line 28)
- methanol only network (43)
- ethanol only network (line 52)
- water-methanol binary network including water network, methanol network, and water-methanol edges (line 61)
- water-ethanol binary network including water network, ethanol network, and water-ethanol edges (line 73)
- methanol-ethanol binary network including methanol network, ethanol network, and methanol-ethanol edges (line 87)

- Entire network: water-methanol-ethanol ternary network including water network, methanol network, ethanol network, water-methanol edges, water-ethanol edges, and methanol-ethanol edges (line 100)

Output Files (./ChemNetworks.exe WaterMethanolEthanol.input water.xyz methanol.xyz ethanol.xyz)
The list of all output file names is shown below.

Water only [GRAPH SOLVENT1 SOLVENT1]

WaterMethanolEthanol.input.water.xyz.water.xyz.Graph

WaterMethanolEthanol.input.water.xyz.water.xyz.GraphGeod

WaterMethanolEthanol.input.water.xyz.water.xyz.GraphNumnodes

Methanol only [GRAPH SOLVENT2 SOLVENT2]

WaterMethanolEthanol.input.methanol.xyz.methanol.xyz.Graph

WaterMethanolEthanol.input.methanol.xyz.methanol.xyz.GraphGeod

WaterMethanolEthanol.input.methanol.xyz.methanol.xyz.GraphNumnodes

Ethanol only [GRAPH SOLVENT3 SOLVENT3]

WaterMethanolEthanol.input.ethanol.xyz.ethanol.xyz.Graph

WaterMethanolEthanol.input.ethanol.xyz.ethanol.xyz.GraphGeod

WaterMethanolEthanol.input.ethanol.xyz.ethanol.xyz.GraphNumnodes

Water-Methanol binary [GRAPH SOLVENT1 SOLVENT2]

WaterMethanolEthanol.input.water.xyz.methanol.xyz.Graph

WaterMethanolEthanol.input.water.xyz.methanol.xyz.GraphGeod

WaterMethanolEthanol.input.water.xyz.methanol.xyz.GraphNumnodes

Water-Ethanol binary [GRAPH SOLVENT1 SOLVENT3]

WaterMethanolEthanol.input.water.xyz.ethanol.xyz.Graph

WaterMethanolEthanol.input.water.xyz.ethanol.xyz.GraphGeod

WaterMethanolEthanol.input.water.xyz.ethanol.xyz.GraphNumnodes

Methanol-Ethanol binary [GRAPH SOLVENT2 SOLVENT3]

WaterMethanolEthanol.input.methanol.xyz.ethanol.xyz.Graph

WaterMethanolEthanol.input.methanol.xyz.ethanol.xyz.GraphGeod

WaterMethanolEthanol.input.methanol.xyz.ethanol.xyz.GraphNumnodes

Water-Methanol-Ethanol ternary [GRAPH SOLVENT1 SOLVENT2 SOLVENT3]

WaterMethanolEthanol.input.water.xyz.methanol.xyz.ethanol.xyz.Graph

WaterMethanolEthanol.input.water.xyz.methanol.xyz.ethanol.xyz.GraphGeod

WaterMethanolEthanol.input.water.xyz.methanol.xyz.ethanol.xyz.GraphNumnodes

A sample output file (for water/ethanol binary network) is given below for illustrating the main features: *WaterMethanolEthanol.input.water.xyz.ethanol.xyz.GraphGeod*

```
2 56 0 0 0
5 165 0 0 0
7 173 0 0 0
12 158 0 0 0
14 29 0 0 0
26 29 0 0 0
32 114 0 0 0        Water-Water
42 175 0 0 0
74 192 0 0 0
79 166 0 0 0
84 111 0 0 0
124 125 0 0 0
142 174 0 0 0
4 97 0 0 1    →  Water-Water PBC (+z image)
230 247 0 0 0
241 245 0 0 0      Ethanol-Ethanol
242 244 0 0 0
8 230 0 0 0
8 247 0 0 0
12 219 0 0 0
38 215 0 0 0
39 240 0 0 0
42 246 0 0 0
46 216 0 0 0
49 233 0 0 0
53 203 0 0 0
63 202 0 0 0
77 202 0 0 0       Water-Ethanol
78 235 0 0 0
84 220 0 0 0
87 217 0 0 0
89 215 0 0 0
105 206 0 0 0
128 201 0 0 0
132 208 0 0 0
143 234 0 0 0
160 225 0 0 0
175 246 0 0 0
183 202 0 0 0
187 250 0 0 0
109 217 0 1 0  →  Water-Ethanol  PBC  (+y image)
188 217 0 1 0
199 245 0 1 0
```

7.0 UTILITIES FOR POST-PROCESSING

Some utility programs will be introduced in this section. A script, written in R language, and a C code will be used to perform degree analysis, network neighborhood analysis, and additional geodesic path analysis.

7.1 R-SCRIPT FOR DEGREE AND NETWORK NEIGHBORHOOD

Once the edge list for the requested graphs/networks are generated using the *ChemNetworks*, the "degree" distributions and histograms can be obtained using our post-processing script written in R language, named *chemical-networks.R,* which reads the *ChemNetworks* output files with the extension .*Graph*. This script also enables to perform the "network neighborhood" analysis. The portions of the script (R language commands) for degree and network neighborhood analyses are shown below:

```
no.net.degs<-degree(grafik)
write(no.net.degs,file="all.degrees",ncolumns=(numbervertices),append=TRUE)
.
.
.
nghbr4<-neighborhood.size(grafik,4,1:(numbervertices))
write(nghbr4,file="neighborhoodsize.order.4",ncolumns=(numbervertices+1))
```

Here, `grafik` represents the graph object, which is essentially the edge-list for the network of interest. The degree (number of edges a vertex is making) of all vertices will be printed in an output file, named `all.degrees`. The fourth order neighborhood size (the number of all the vertices within 4 order neighborhood) of all vertices will be printed in an output file, named `neighborhoodsize.order.4`.

The script *chemical-networks.R* is provided along with the *ChemNetworks* package. To run this script, the *igraph* R library should also be installed from the CRAN website: http://cran.r-project.org.

7.2 CODE FOR GEODESIC PATH ANALYSIS

Once the geodesic distance matrix is obtained from the *ChemNetworks*, further analyses of the geodesics paths of each length can be carried out using the utility code, *geodesic-statistics.c.* This C code reads the *ChemNetworks* output files with the extension *.geocard*. The file *geodesic-statistics.c* is provided along with the *ChemNetworks* package. It can be used to obtain a detailed histogram of the geodesics for the network of interst cross-correlated with the Euclidean distance between the terminal vertices of each geodesic path. Moreover, *geodesic-statistics.c* can be use to determine the persistence/lifetime of the specific length geodecis paths.

8.0 HOW TO CITE *ChemNetworks*

In publishing results obtained either in part or in full from use of *ChemNetworks*, the user should use the following citation:

Citation:

Ozkanlar, A; Clark, A. E. "*ChemNetworks*: A Complex Network Analysis Tool For Chemical Systems" *J. Comp. Chem.* **2013,** *Volume*, pages.