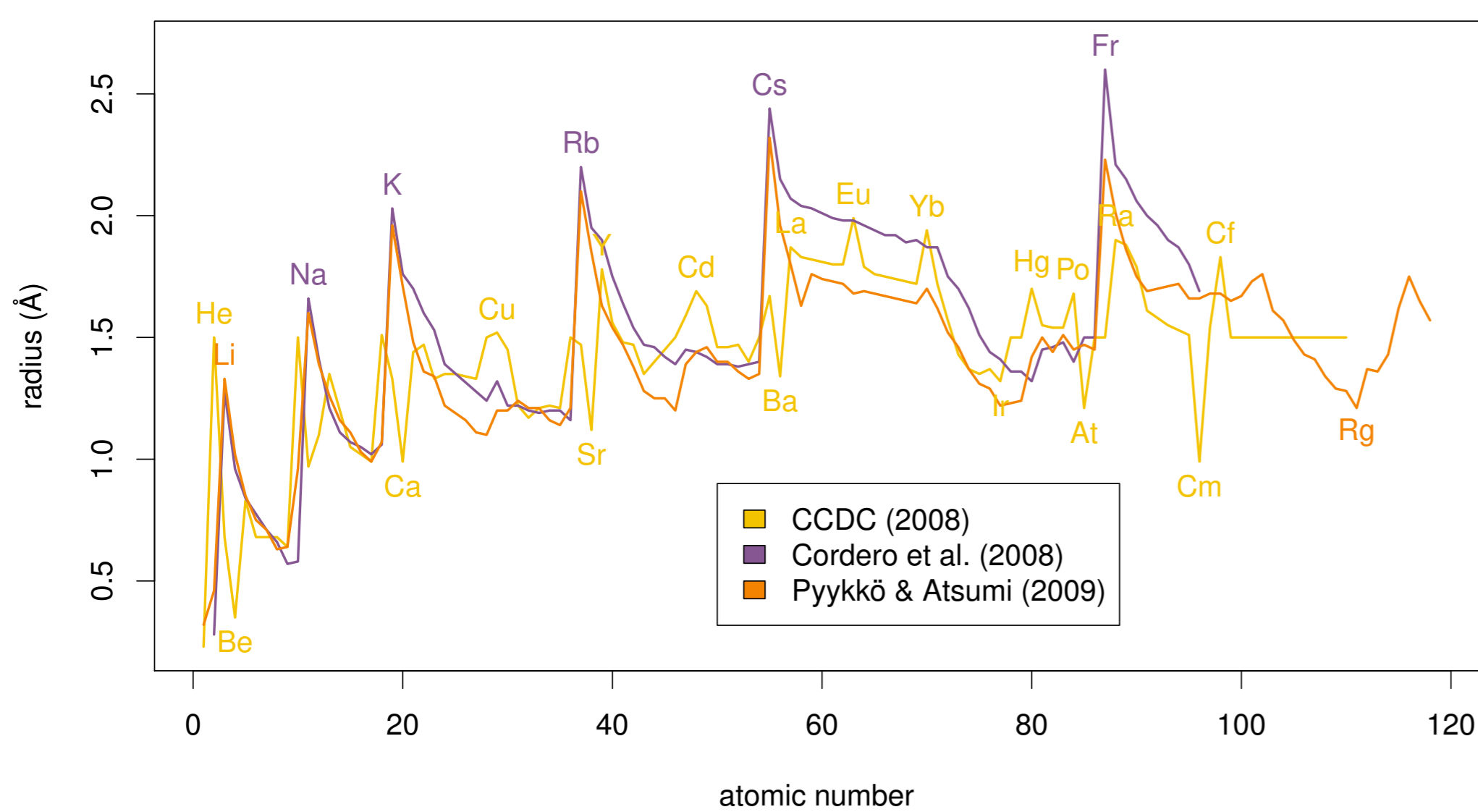




## Abstract

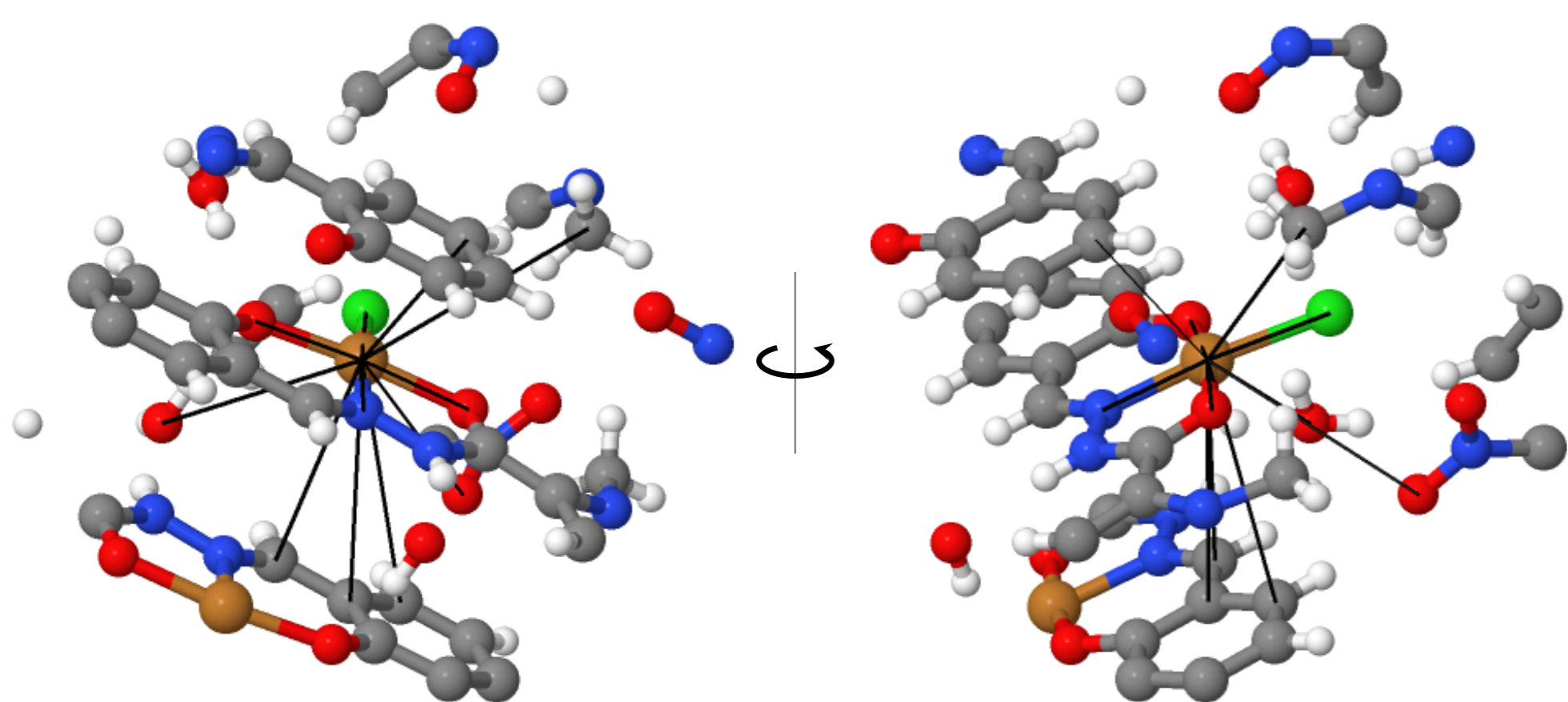
Methods of crystal structure determination, most prominent of them being X-ray crystallography, provide exact 3D coordinates of the atoms that constitute crystal structures. Chemical bonding, which is usually of interest to the chemists, is not captured by crystal structure determination methods and has to be inferred from the coordinates of the atoms. Usually, two atoms are assumed to be connected by a chemical bond if the distance between them is shorter than the sum of their element-specific bonding radii, typically covalent radii, that is an usual radius of an atom in a covalent bond. This method has been known at least since the publications of William Lawrence Bragg [1]. However, there is no one univocal table of covalent radii as the methodologies for defining and deriving them differ strongly. It has been noted that clear distinctions exist in the distance distributions of atom pairs, allowing to differentiate between bonded and non-bonded interactions. Unpopulated ranges correspond to the so-called van der Waals gap, a region of energetically unfavourable interatomic distances, which may in some cases be contaminated due to the presence of non-covalent interactions. Furthermore, distance distributions of some elements, such as alkaline elements, copper, silver, mercury, iron, tin to name a few, do not have such a clear van der Waals gap and pose a difficulty in fitting them to the aforementioned approach [2]. Nevertheless, distance criterion for connectivity is widely used, and covalent radii for problematic elements have been approximated. We have made an attempt to investigate the bonding radii of atoms using a statistical approach. Distances between nearest neighbouring atoms, as established by 3D Voronoi tessellation of the crystal [3], were measured in each of the 380 000 crystal structures in the Crystallography Open Database (COD, <http://www.crystallography.net>) [4] and pooled together for analysis. Limiting observations to those of the neighbouring cells of Voronoi tessellation allowed to curb the most of the non-informative random distribution of long distances without the need of artificial cut-offs. Visual analysis of the resulting histograms reveal clearly separated high density regions corresponding to the covalent and van der Waals interactions as well as other features with chemical meaning. To identify these features, automatic unsupervised statistical methods [5] could be harnessed.

## Covalent radii tables



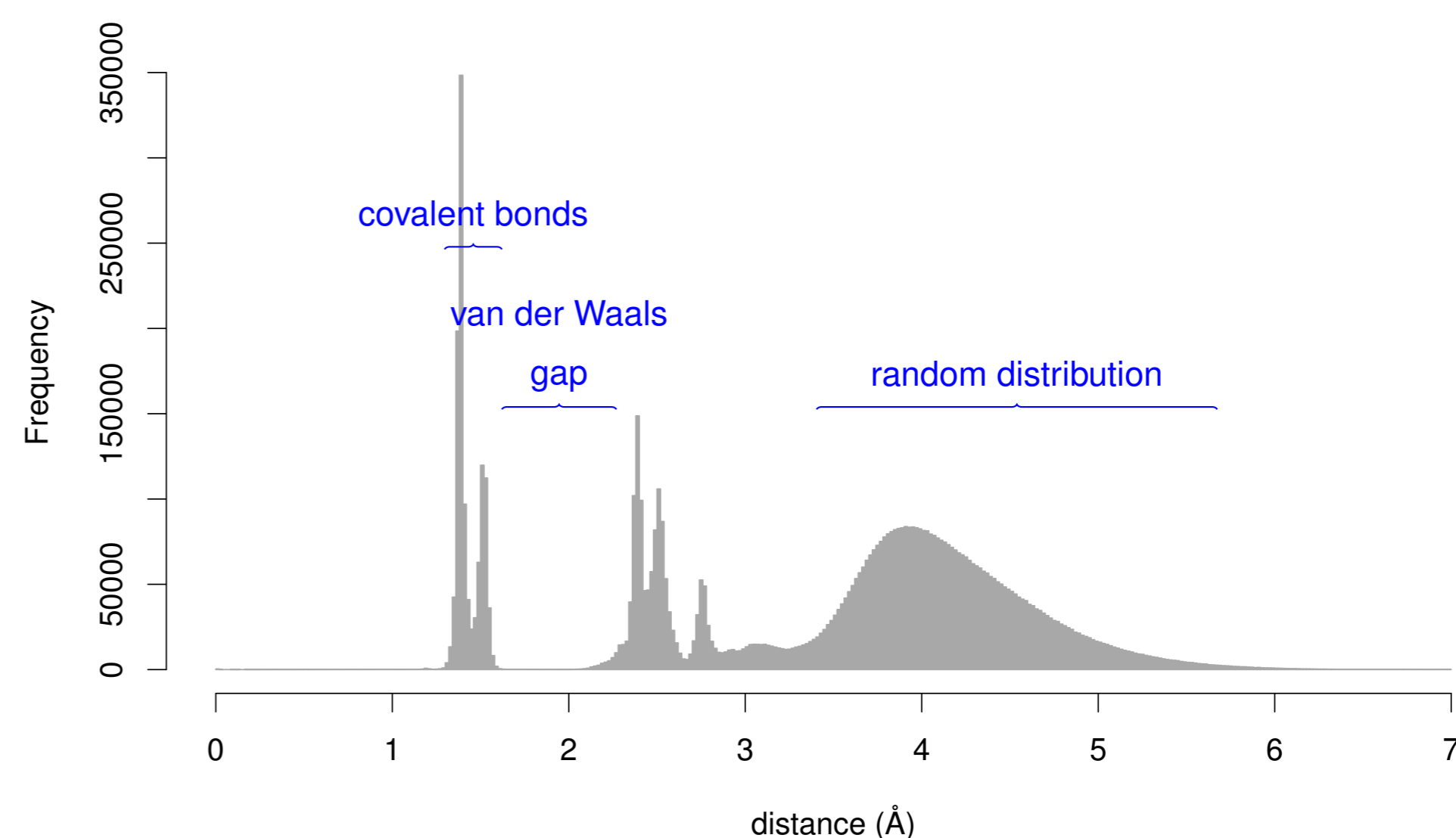
- ▶ Differ vastly;
- ▶ Depend on used assumptions and target applications;
- ▶ Usually derived by visual inspection and use artificial cut-offs.

## Contact discovery based on Voronoi tessellation



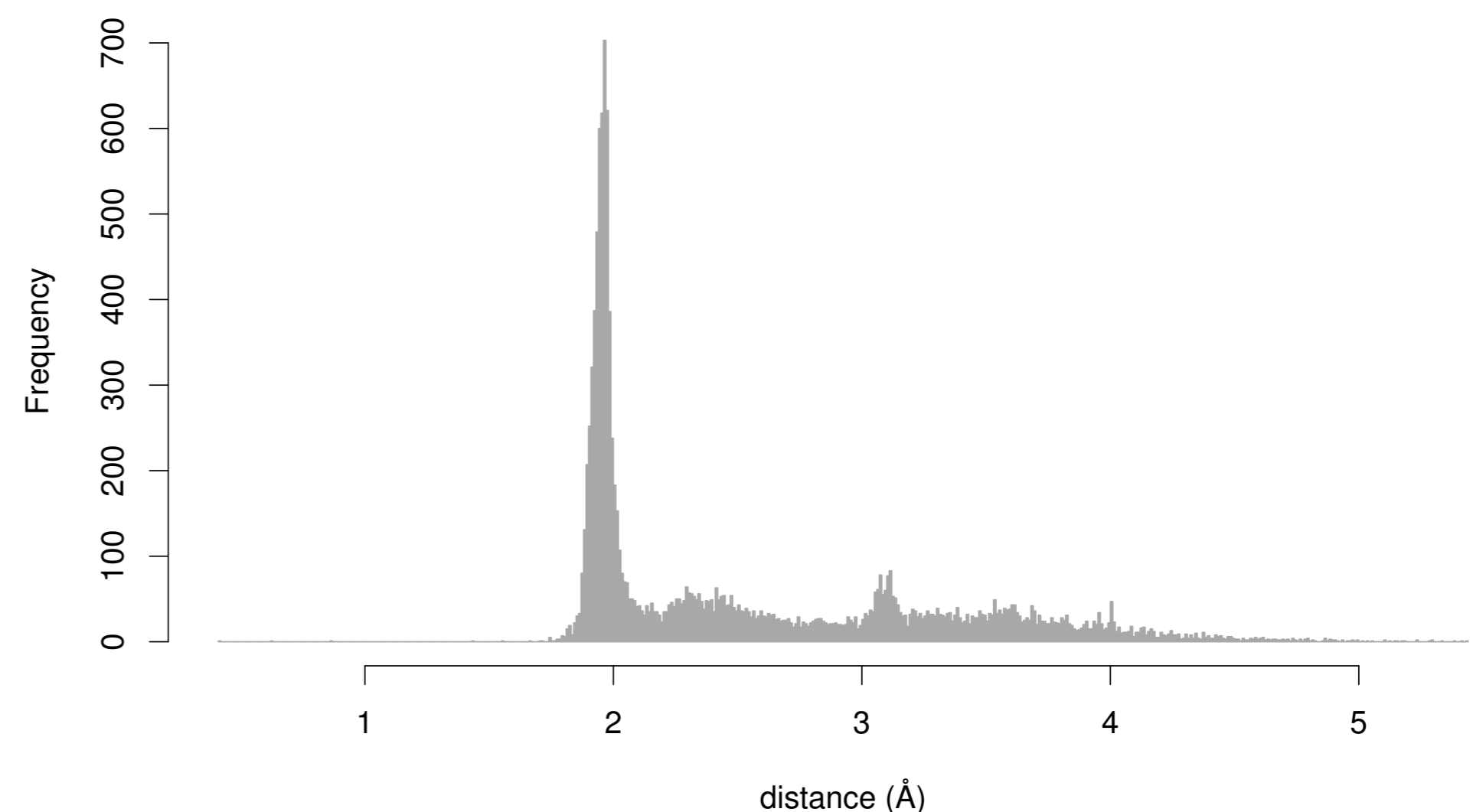
- ▶ Efficiently finds the first coordination sphere for each atom;
- ▶ No cut-offs are required;
- ▶ Hydrogen atoms are ignored due to less reliable coordinates.

## Carbon–carbon distances in crystal structures



- ▶ Unfavourable distances are easily identifiable;
- ▶ Clear distinction between bonded and non-bonded interactions.

## Copper–oxygen distances in crystal structures



- ▶ Separation of bonded and non-bonded interactions is not trivial;
- ▶ Influence of Jahn–Teller effect in six-coordinate Cu(II) complexes is evident as reported before [6].

## Challenges

- ▶ Crystal voids due to missing solvent molecules;
- ▶ Disordered parts of crystals;
- ▶ Big data ( $N$  atoms yield up to  $N^2$  interactions);
- ▶ Automatic unsupervised data analysis.

## Bibliography

- [1] Cordero et al. Covalent radii revisited. *Dalton Transactions*, pages 2832–2838, 2008.
- [2] Alvarez. A cartography of the van der Waals territories. *Dalton Transactions*, 42:8617–8636, 2013.
- [3] Olechnovič et al. Voronota: A fast and reliable tool for computing the vertices of the Voronoi diagram of atomic balls. *Journal of Computational Chemistry*, 35:672–681, 2014.
- [4] Gražulis et al. Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Research*, 40:D420–D427, 2012.
- [5] Merkys et al. Stereochemical statistics in Crystallography Open Database. *Acta Crystallographica Section A*, 69:s388–s389, 2013.
- [6] Harding. The geometry of metal-ligand interactions relevant to proteins. *Acta Crystallographica Section D*, 55:1432–1443, 1999.
- [7] The Cambridge Crystallographic Data Centre. Element data and radii, 2008. <https://web.archive.org/web/20080701015237/http://www.ccdc.c...>
- [8] Pyykkö et al. Molecular single-bond covalent radii for elements 1–118. *Chemistry – A European Journal*, 15:186–197, 2009.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 689868.

On-line version of the poster:  
<http://bit.ly/2DEk6HL>

