Meeting 1: Enumeration and Heuristics

Stefan Gugler

Massachusetts Institute of Technology

sgugler@mit.edu

March 28, 2018

Mono-Heavy-Atomic Isoelectronic Ligand Distribution

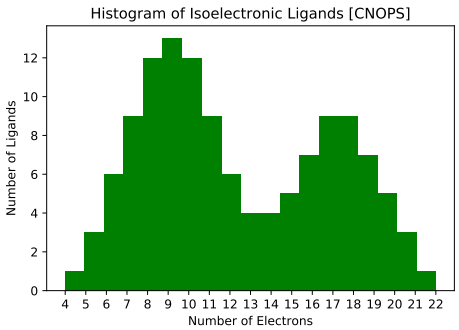

Figure: CNOPS: 15625 ligands

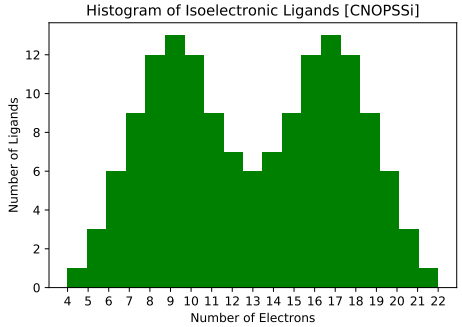

Figure: CNOPSSi: 22500 (?) ligands

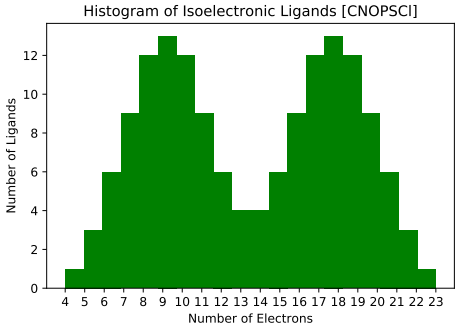

Figure: CNOPSCl: 22500 (?) ligands

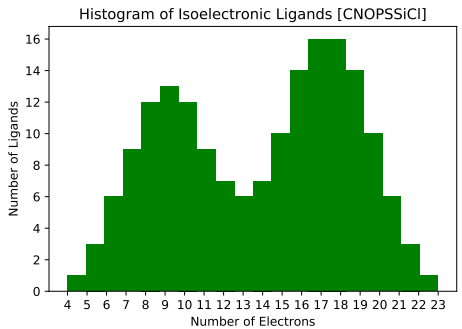

Figure: CNOPSSiCl: 30625 ligands

Mono-Heavy-Atomic Iso-VE Ligand Distribution

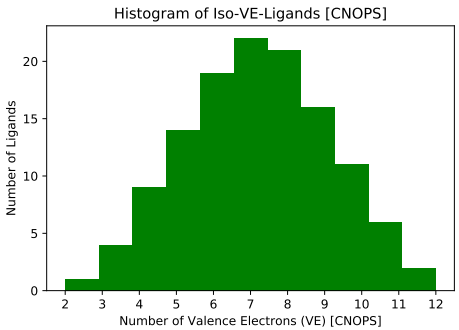

Figure: CNOPS: 15625 ligands

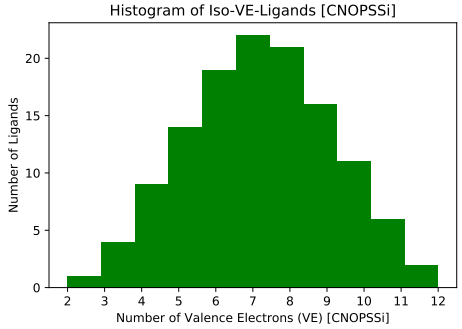

Figure: CNOPSSi: 22500 (?) ligands

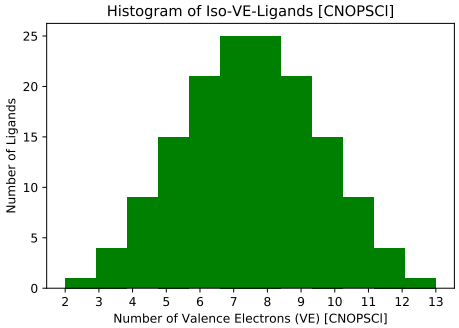

Figure: CNOPSCl: 22500 (?) ligands

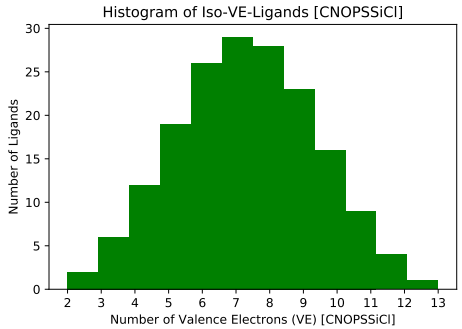

Figure: CNOPSSiCl: 30625 ligands

Now we go to Di-Heavy-Atomic Ligands ...

- Exhaustive combinatorial combination from the Mono-Heavy-Atomic Ligands to get Di-Heavy-Ligands
- Over heavy atoms identities $a_i \in \{\mathrm{C}, \mathrm{N}, \mathrm{O}, \mathrm{P}, \mathrm{S}, (\mathrm{Si}, \mathrm{Cl})\}$, charges $c_i \in \{-2, +2\}$, number of H atoms $h_i \in \{0, 4\}$, with heavy atoms $i \in 1, 2$, where $i = 1$ is the coordinating atom.
- Pruned by means of a health function, u, of ligand j, to encoding heuristics.

Now we go to Di-Heavy-Atomic Ligands ...

- Only ligands with $c_1 \leq 0 \wedge c_2 \leq 0$ are considered.

$$u_{\text{octet,i}} = \begin{cases} 10 + 2 \cdot (8 - VE_i) & \text{if } 8 - VE_i < 0 \\ 10 - 1 \cdot (8 - VE_i) & \text{if } 8 - VE_i \geq 0 \end{cases} \tag{1}$$

$$u_{\text{charge}} = \begin{cases} 0 & \text{if } c_1 + c_2 > 0 \\ 3 & \text{if } 0 \geq c_1 + c_2 \geq -2 \\ 1 & \text{if } c_1 + c_2 = -3 \\ 0 & \text{if } c_1 + c_2 = -4 \end{cases} \tag{2}$$

$$u_{\text{VSEPR}} = 5 - |\underbrace{(VE_1 - 2 \cdot LP_1 + c_1 - 2 \cdot h_1)}_{\textit{ready electrons}} - \underbrace{(VE_2 - 2 \cdot LP_2 + c_2 - 2 \cdot h_2)}_{\textit{ready electrons}}| \tag{3}$$

Now we go to Di-Heavy-Atomic Ligands ...

$$u_{\mathrm{CA}} = \begin{cases} 1 & \text{if } h_1 = 4 \\ 2 & \text{if } h_1 = 3 \\ 3 & \text{if else} \end{cases} \tag{4}$$

$$u_{\mathrm{shell}} = \begin{cases} 1 & \text{if } \neg(VE_1 + VE_2)\%2 \\ 0 & \text{if else} \end{cases} \tag{5}$$

$$u_{\mathrm{total}} = u_{\mathrm{shell}} \cdot \left(\frac{1}{2}(u_{\mathrm{octet},1} + u_{\mathrm{octet,i}}) + u_{\mathrm{charge}} + u_{\mathrm{VSEPR}} + u_{\mathrm{CA}} \right) \tag{6}$$

where VE_i denotes the number of valence electrons, LP_i the number of lone pairs.

Distribution of the Score

Comparison of edge cases. Mayer B.O. with `multiwfn`.

1	2	Mayer B.O.	score	comment
CHHHH++	CHHHH++	<0.05	12	rather high
CHHHH–	CHHHH–	0.928	-	$c_{tot} > 0$
CHHH	NH	1.308	0	open shell
CH4++	-	0.577	-	Mono-Heavy-Atom
C-	O+	-	16	dummy

Distribution of the Score

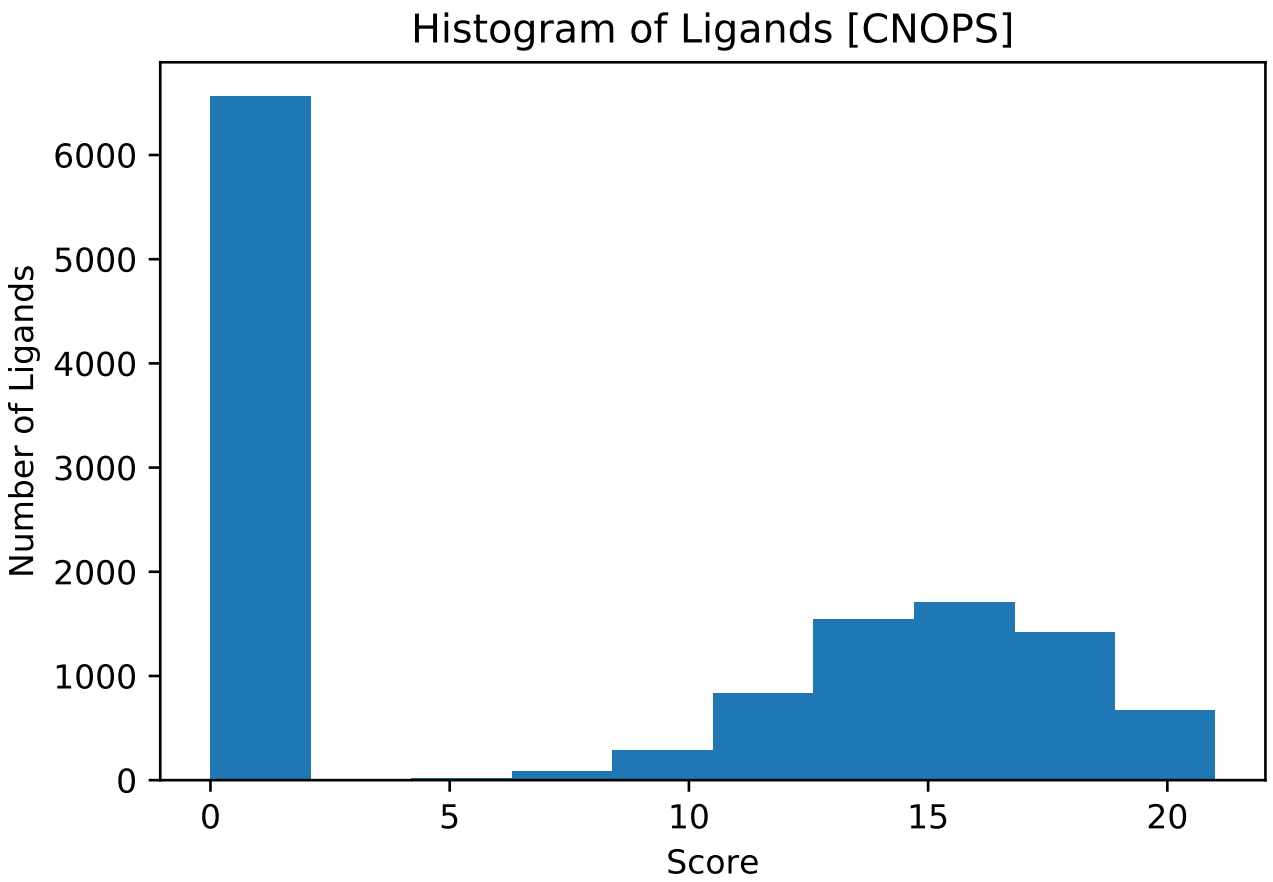

Figure: A histogram describing the score distribution. There are 30 ligands with score 21 and 180 ligands with score 20. CO is has a score of 16. There are 2917 ligands at least as good as CO.

Code Output

```
[['[NHH-]', 10, 8, 1, -1, 2, 0, 1, 10, '[CHHH-]', 10, 8, 0, -1, 3, 0, 1, 10], 21.0]
['[NHH-', 10, 8, 1, -1, 2, 0, 1, 10, '[NHH-]', 10, 8, 1, -1, 2, 0, 1, 10], 21.0]
['[NHH-', 10, 8, 1, -1, 2, 0, 1, 10, '[OH-]', 10, 8, 2, -1, 1, 0, 1, 10], 21.0]
['[NHH-', 10, 8, 1, -1, 2, 0, 1, 10, '[PHH-]', 18, 8, 1, -1, 2, 0, 1, 10], 21.0]
['[NHH-', 10, 8, 1, -1, 2, 0, 1, 10, '[SH-]', 18, 8, 2, -1, 1, 0, 1, 10], 21.0]
['[OH-', 10, 8, 2, -1, 1, 0, 1, 10, '[CHHH-]', 10, 8, 0, -1, 3, 0, 1, 10], 21.0]
['[OHH', 10, 8, 2, 0, 2, 0, 1, 10, '[CHHHH]', 10, 8, 0, 0, 4, 0, 1, 10], 21.0]
['[OH-', 10, 8, 2, -1, 1, 0, 1, 10, '[NHH-]', 10, 8, 1, -1, 2, 0, 1, 10], 21.0]
['[OHH', 10, 8, 2, 0, 2, 0, 1, 10, '[NHHH]', 10, 8, 1, 0, 3, 0, 1, 10], 21.0]
['[OH-', 10, 8, 2, -1, 1, 0, 1, 10, '[OH-]', 10, 8, 2, -1, 1, 0, 1, 10], 21.0]
['[OHH', 10, 8, 2, 0, 2, 0, 1, 10, '[OHH]', 10, 8, 2, 0, 2, 0, 1, 10], 21.0]
```