



# How many materials are left to discover?

*An exploration of quaternary space*

Presentation by **Michael Sluydts**

Research performed by **Michael Sluydts, Michiel Larmuseau, Karel Dumon, Titus Crepain, Kurt Lejaeghere and Stefaan Cottenier**

**Center for Molecular Modeling  
Ghent University  
Belgium**



# Finding the right material

- 1) How many materials are there left to discover?
- 2) How do we find them?



Can we create a Drake equation for materials?

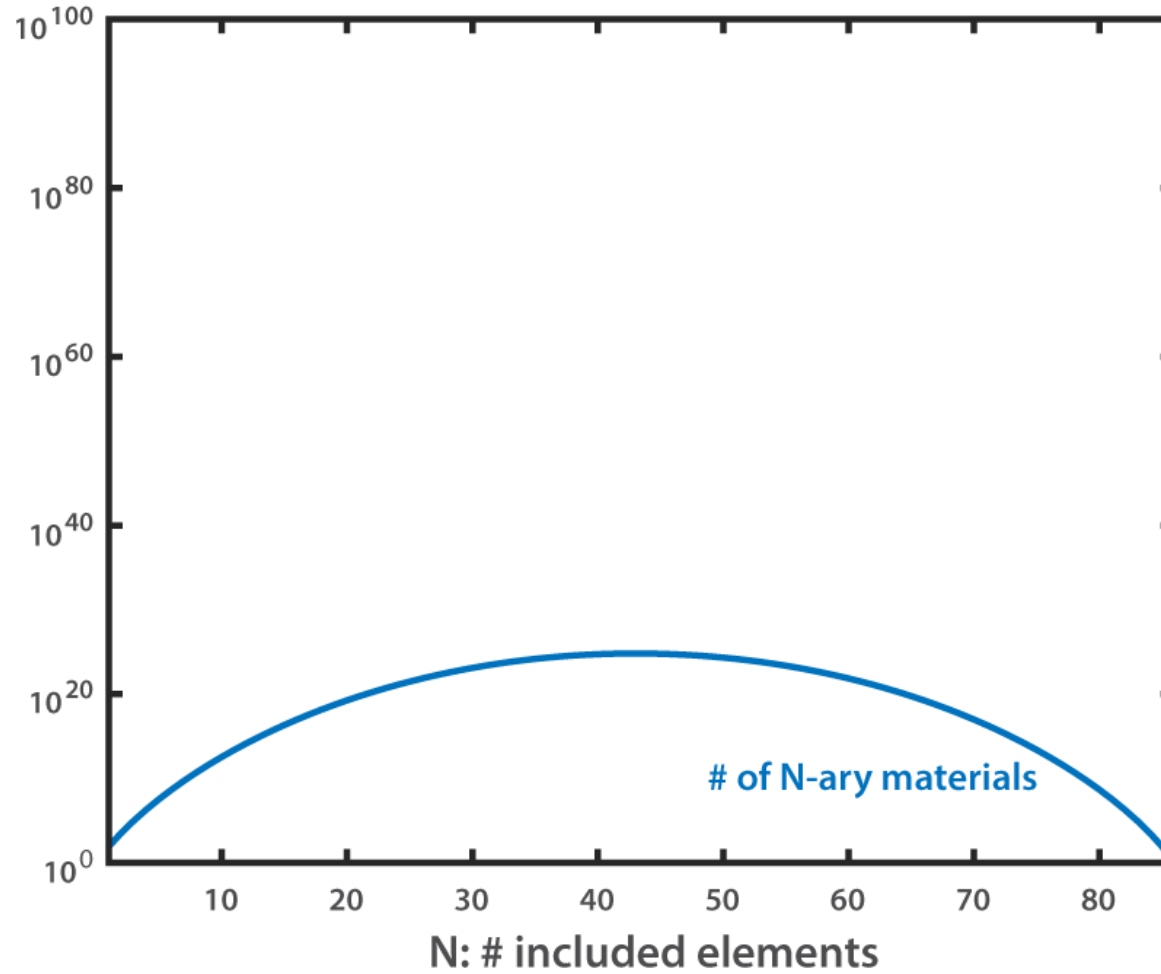
# Periodic table

hydrogen 1 <b>H</b> 1.0079																	helium 2 <b>He</b> 4.0026									
lithium 3 <b>Li</b> 6.941	beryllium 4 <b>Be</b> 9.0122											boron 5 <b>B</b> 10.811	carbon 6 <b>C</b> 12.011	nitrogen 7 <b>N</b> 14.007	oxygen 8 <b>O</b> 15.999	fluorine 9 <b>F</b> 18.998	neon 10 <b>Ne</b> 20.180									
sodium 11 <b>Na</b> 22.990	magnesium 12 <b>Mg</b> 24.305											aluminum 13 <b>Al</b> 26.982	silicon 14 <b>Si</b> 28.086	phosphorus 15 <b>P</b> 30.974	sulfur 16 <b>S</b> 32.065	chlorine 17 <b>Cl</b> 35.453	argon 18 <b>Ar</b> 39.948									
potassium 19 <b>K</b> 39.098	calcium 20 <b>Ca</b> 40.078	scandium 21 <b>Sc</b> 44.956	titanium 22 <b>Ti</b> 47.867	vanadium 23 <b>V</b> 50.942	chromium 24 <b>Cr</b> 51.996	manganese 25 <b>Mn</b> 54.938	iron 26 <b>Fe</b> 55.845	cobalt 27 <b>Co</b> 58.933	nickel 28 <b>Ni</b> 58.693	copper 29 <b>Cu</b> 63.546	zinc 30 <b>Zn</b> 65.38	gallium 31 <b>Ga</b> 69.723	germanium 32 <b>Ge</b> 72.64	arsenic 33 <b>As</b> 74.922	selenium 34 <b>Se</b> 78.96	bromine 35 <b>Br</b> 79.904	krypton 36 <b>Kr</b> 83.798									
rubidium 37 <b>Rb</b> 85.468	strontium 38 <b>Sr</b> 87.62	yttrium 39 <b>Y</b> 88.906	zirconium 40 <b>Zr</b> 91.224	niobium 41 <b>Nb</b> 92.906	molybdenum 42 <b>Mo</b> 95.96	technetium 43 <b>Tc</b> [98]	ruthenium 44 <b>Ru</b> 101.07	rhodium 45 <b>Rh</b> 102.91	palladium 46 <b>Pd</b> 106.42	silver 47 <b>Ag</b> 107.87	cadmium 48 <b>Cd</b> 112.41	indium 49 <b>In</b> 114.82	tin 50 <b>Sn</b> 118.71	antimony 51 <b>Sb</b> 121.76	tellurium 52 <b>Te</b> 127.60	iodine 53 <b>I</b> 126.90	xenon 54 <b>Xe</b> 131.29									
cesium 55 <b>Cs</b> 132.91	barium 56 <b>Ba</b> 137.33											hafnium 72 <b>Hf</b> 178.49	tantalum 73 <b>Ta</b> 180.95	tungsten 74 <b>W</b> 183.84	rhenium 75 <b>Re</b> 186.21	osmium 76 <b>Os</b> 190.23	iridium 77 <b>Ir</b> 192.22	platinum 78 <b>Pt</b> 195.08	gold 79 <b>Au</b> 196.97	mercury 80 <b>Hg</b> 200.59	thallium 81 <b>Tl</b> 204.38	lead 82 <b>Pb</b> 207.2	bismuth 83 <b>Bi</b> 208.98	polonium 84 <b>Po</b> [209]	astatine 85 <b>At</b> [210]	radon 86 <b>Rn</b> [222]
francium 87 <b>Fr</b> [223]	radium 88 <b>Ra</b> [226]	rutherfordium 104 <b>Rf</b> [261]	dubnium 105 <b>Db</b> [262]	seaborgium 106 <b>Sg</b> [266]	bohrium 107 <b>Bh</b> [264]	hassium 108 <b>Hs</b> [277]	meitnerium 109 <b>Mt</b> [268]	darmstadtium 110 <b>Ds</b> [271]	roentgenium 111 <b>Rg</b> [272]																	

lanthanum 57 <b>La</b> 138.91	cerium 58 <b>Ce</b> 140.12	praseodymium 59 <b>Pr</b> 140.91	neodymium 60 <b>Nd</b> 144.24	promethium 61 <b>Pm</b> [145]	smarium 62 <b>Sm</b> 150.36	europium 63 <b>Eu</b> 151.96	gadolinium 64 <b>Gd</b> 157.25	terbium 65 <b>Tb</b> 158.93	dysprosium 66 <b>Dy</b> 162.50	holmium 67 <b>Ho</b> 164.93	erbium 68 <b>Er</b> 167.26	thulium 69 <b>Tm</b> 168.93	ytterbium 70 <b>Yb</b> 173.05	lutetium 71 <b>Lu</b> 174.97
actinium 89 <b>Ac</b> [227]	thorium 90 <b>Th</b> 232.04	protactinium 91 <b>Pa</b> 231.04	uranium 92 <b>U</b> 238.03	neptunium 93 <b>Np</b> [237]	plutonium 94 <b>Pu</b> [244]	americium 95 <b>Am</b> [243]	curium 96 <b>Cm</b> [247]	berkelium 97 <b>Bk</b> [247]	californium 98 <b>Cf</b> [251]	einsteinium 99 <b>Es</b> [252]	fermium 100 <b>Fm</b> [257]	mendelevium 101 <b>Md</b> [258]	nobelium 102 <b>No</b> [259]	lawrencium 103 <b>Lr</b> [262]

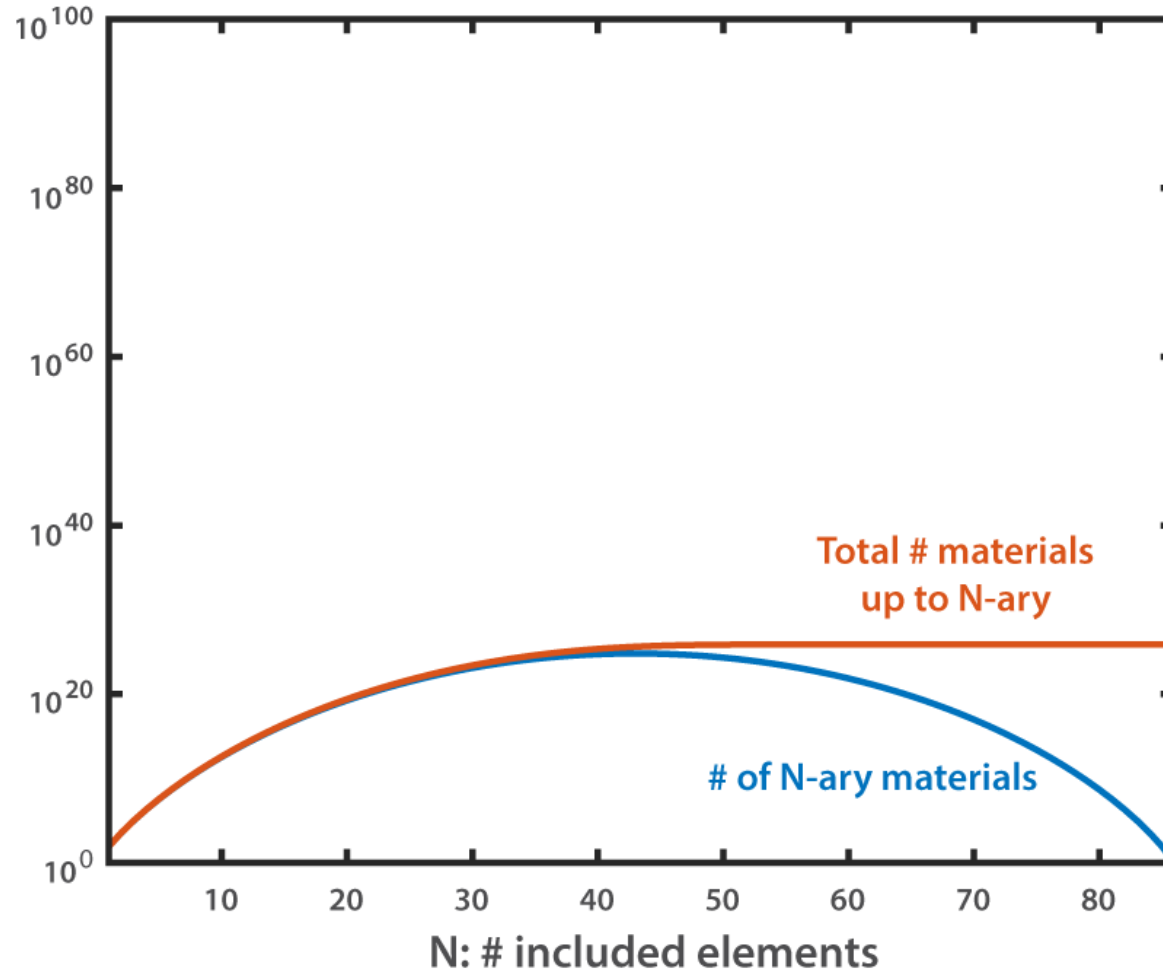
Consider the first 86 elements in the periodic table

# All combinations



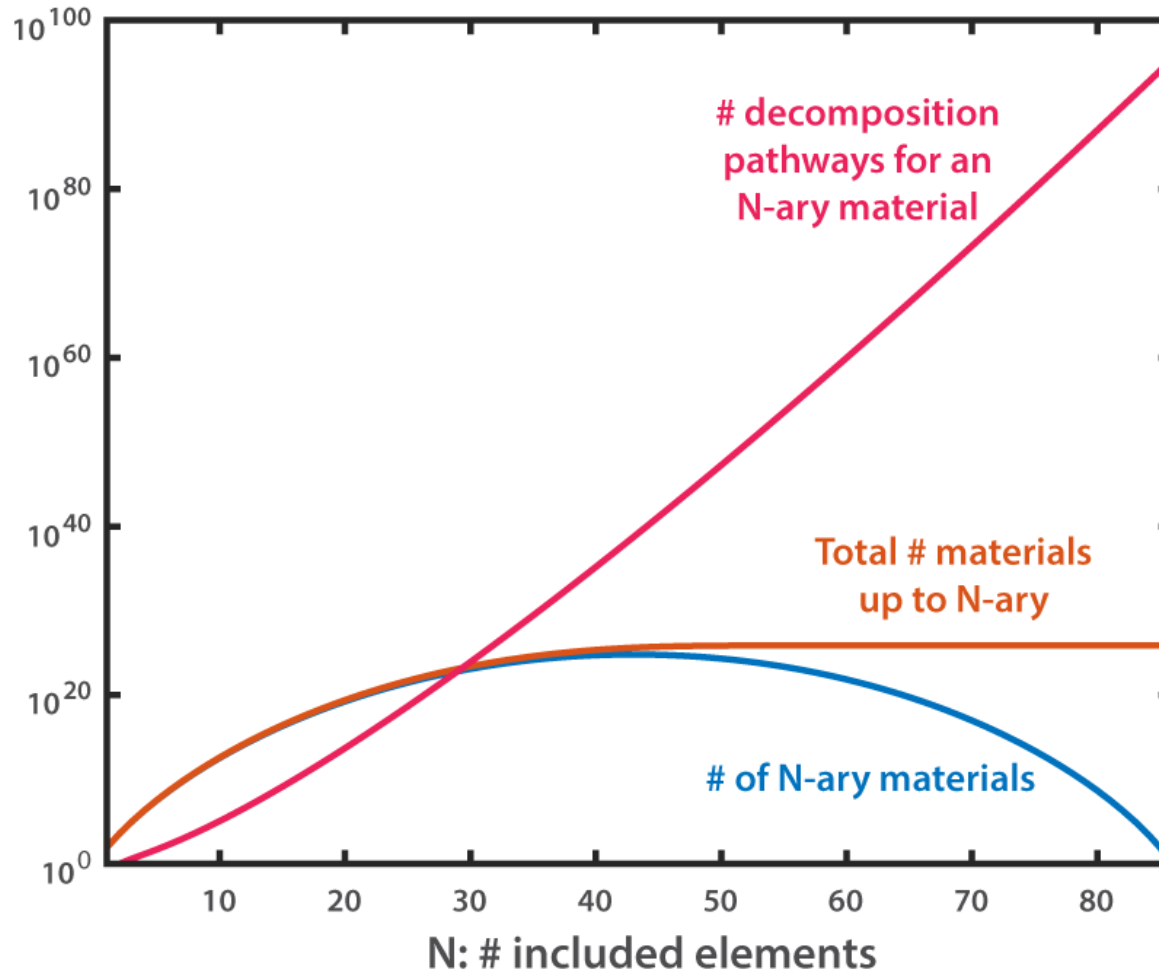
**Maximum at 43 elements:  $10^{25}$  materials**

# All combinations



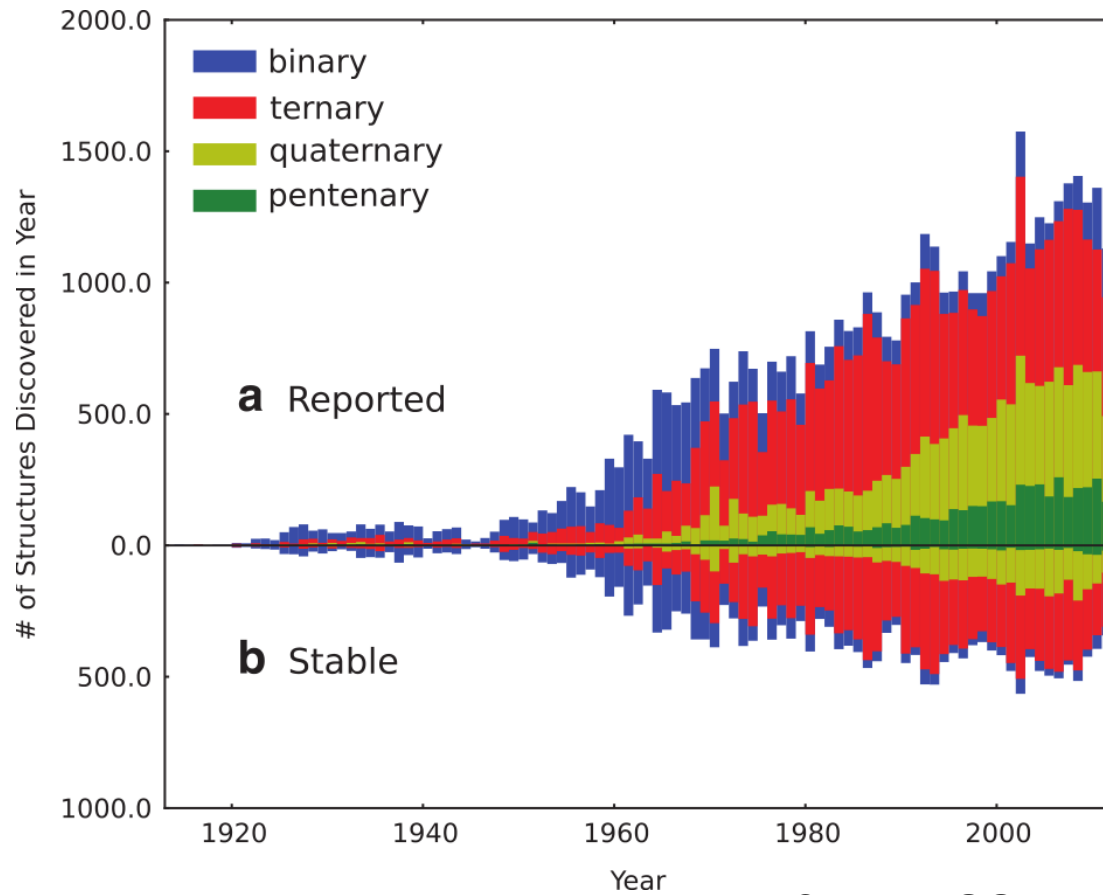
**Total dominated by maximum**

# Decomposition pathways



**Decomposition pathways increase even faster**

# The current status



Source: OQMD

Only about 500 quaternaries are found each year

# QZP space

Quaternary zintl phases provide an interesting subset of materials

hydrogen 1 <b>H</b> 1.0079																								helium 2 <b>He</b> 4.0026
lithium 3 <b>Li</b> 6.941	beryllium 4 <b>Be</b> 9.0122											boron 5 <b>B</b> 10.811	carbon 6 <b>C</b> 12.011	nitrogen 7 <b>N</b> 14.007	oxygen 8 <b>O</b> 15.999	fluorine 9 <b>F</b> 18.998	neon 10 <b>Ne</b> 20.180							
sodium 11 <b>Na</b> 22.990	magnesium 12 <b>Mg</b> 24.305											aluminum 13 <b>Al</b> 26.982	silicon 14 <b>Si</b> 28.086	phosphorus 15 <b>P</b> 30.974	sulfur 16 <b>S</b> 32.065	chlorine 17 <b>Cl</b> 35.453	argon 18 <b>Ar</b> 39.948							
potassium 19 <b>K</b> 39.098	calcium 20 <b>Ca</b> 40.078	scandium 21 <b>Sc</b> 44.956	titanium 22 <b>Ti</b> 47.867	vanadium 23 <b>V</b> 50.942	chromium 24 <b>Cr</b> 51.996	manganese 25 <b>Mn</b> 54.938	iron 26 <b>Fe</b> 55.845	cobalt 27 <b>Co</b> 58.933	nickel 28 <b>Ni</b> 58.693	copper 29 <b>Cu</b> 63.546	zinc 30 <b>Zn</b> 65.38	gallium 31 <b>Ga</b> 69.723	germanium 32 <b>Ge</b> 72.64	arsenic 33 <b>As</b> 74.922	selenium 34 <b>Se</b> 78.96	bromine 35 <b>Br</b> 79.904	krypton 36 <b>Kr</b> 83.798							
rubidium 37 <b>Rb</b> 85.468	strontium 38 <b>Sr</b> 87.62	yttrium 39 <b>Y</b> 88.906	zirconium 40 <b>Zr</b> 91.224	niobium 41 <b>Nb</b> 92.906	molybdenum 42 <b>Mo</b> 95.96	technetium 43 <b>Tc</b> [98]	ruthenium 44 <b>Ru</b> 101.07	rhodium 45 <b>Rh</b> 102.91	palladium 46 <b>Pd</b> 106.42	silver 47 <b>Ag</b> 107.87	cadmium 48 <b>Cd</b> 112.41	indium 49 <b>In</b> 114.82	tin 50 <b>Sn</b> 118.71	antimony 51 <b>Sb</b> 121.76	tellurium 52 <b>Te</b> 127.60	iodine 53 <b>I</b> 126.90	xenon 54 <b>Xe</b> 131.29							
caesium 55 <b>Cs</b> 132.91	barium 56 <b>Ba</b> 137.33	lanthanum 57 <b>La</b> 138.91	hafnium 72 <b>Hf</b> 178.49	tantalum 73 <b>Ta</b> 180.95	tungsten 74 <b>W</b> 183.84	rhenium 75 <b>Re</b> 186.21	osmium 76 <b>Os</b> 190.23	iridium 77 <b>Ir</b> 192.22	platinum 78 <b>Pt</b> 195.08	gold 79 <b>Au</b> 196.97	mercury 80 <b>Hg</b> 200.59	thallium 81 <b>Tl</b> 204.38	lead 82 <b>Pb</b> 207.2	bismuth 83 <b>Bi</b> 208.98	polonium 84 <b>Po</b> [209]	astatine 85 <b>At</b> [210]	radon 86 <b>Rn</b> [222]							

Zintl phases combine covalent and ionic bonding in a single crystal



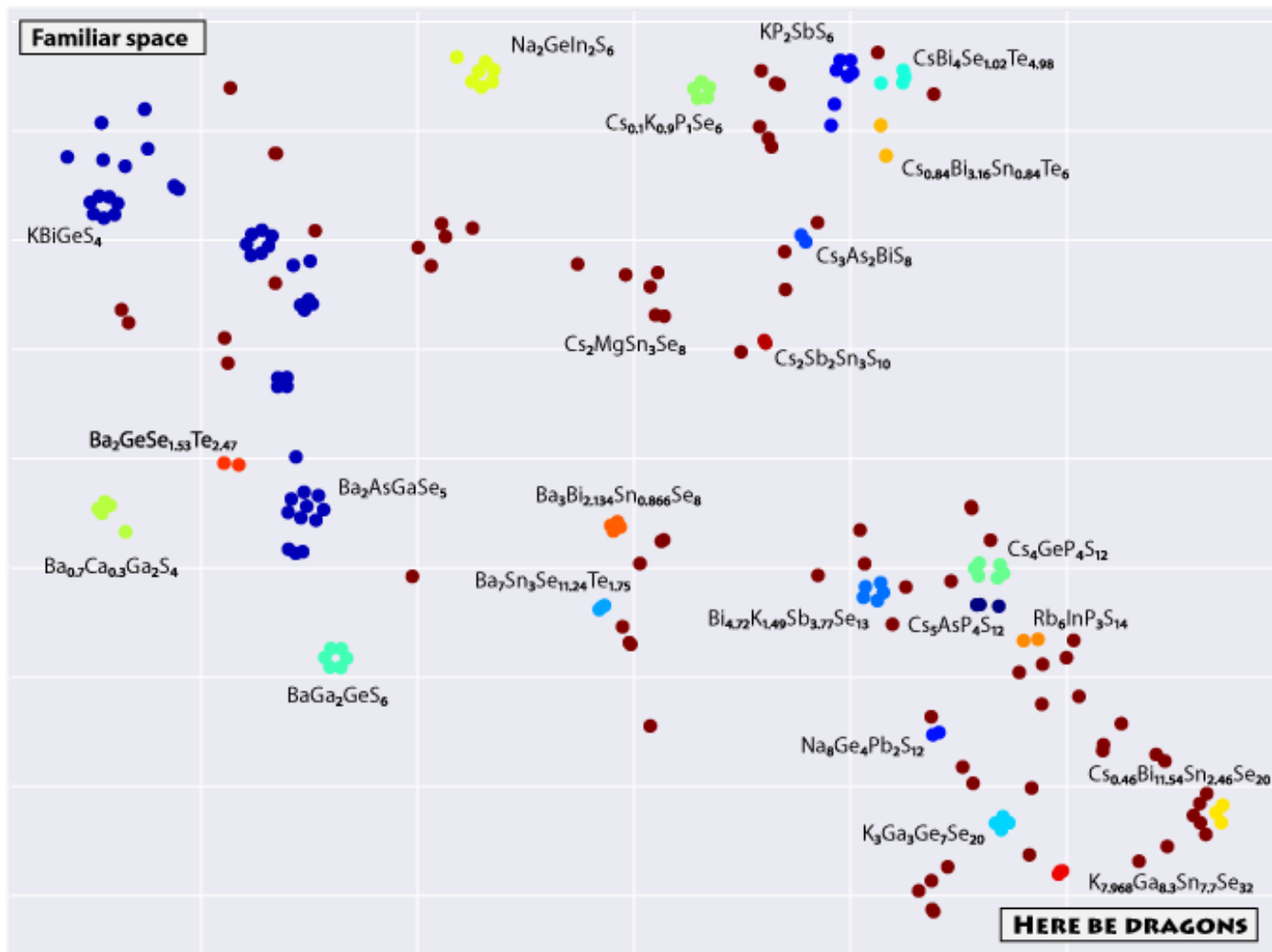
# Visualizing QZP space

Let's create a map of QZP space:

- Extract all known structures from ICSD
- Discard partial occupancies (around 2/3)
- Generate features:  

atomic+ stoichiometry + spacegroup
- Cluster with DBSCAN
- Visualize with t-SNE

# Visualizing QZP space



Similar structures are clustered, but many are unique.

# Drake for quaternaries

Estimating the number of materials based on observed QZP is like estimating the number of planets based on the milky way

## Simplified Drake equation

$N_{planets\ with\ life}$

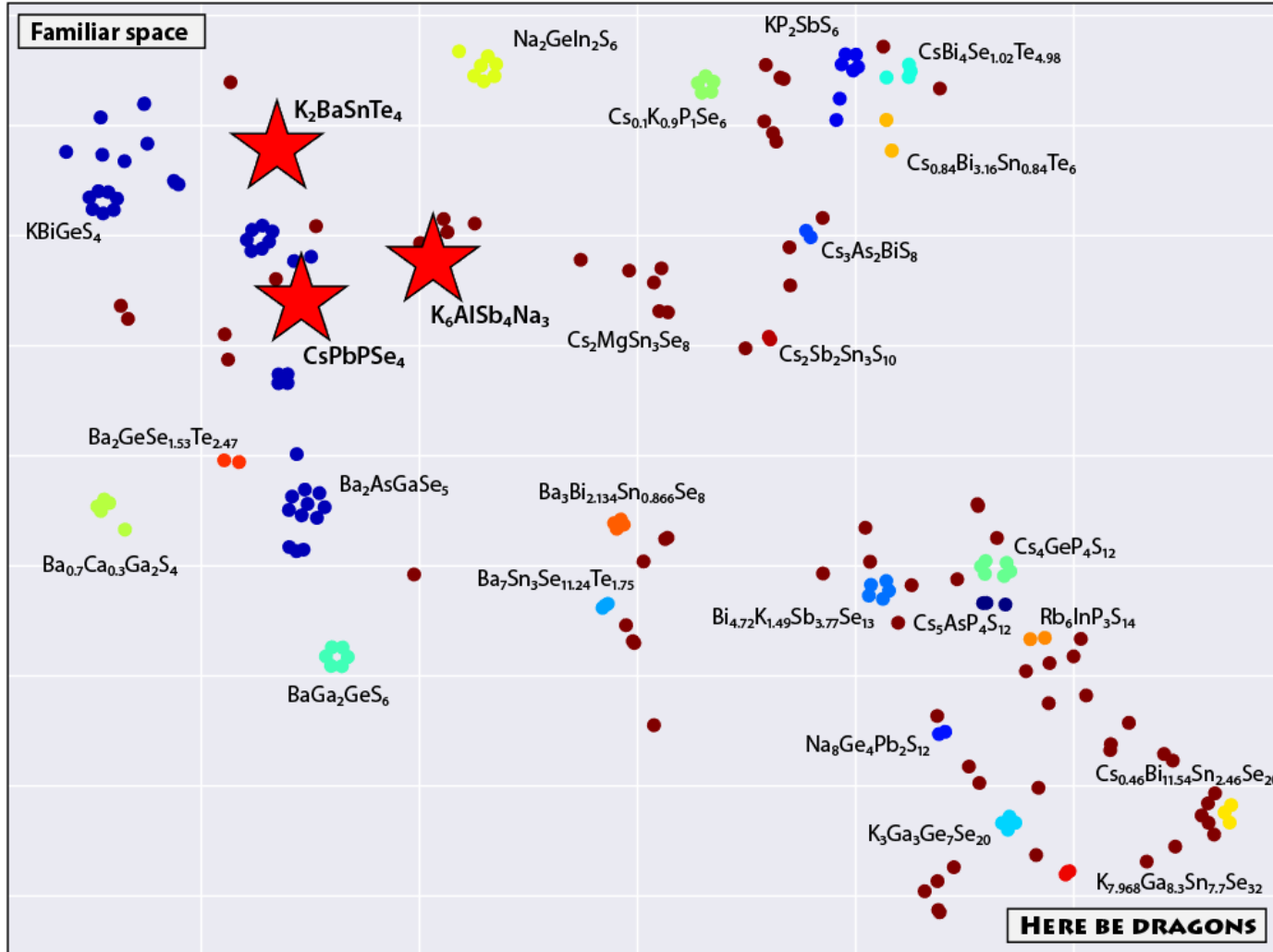
$$= \frac{N_{planets,universe}}{N_{planets,MW}} \cdot N_{solar\ systems,MW} \cdot N_{\frac{observed}{solar\ system}} \cdot N_{\frac{total}{observed}} \cdot f_{life}$$

## Materials Drake equation

$N_{useful\ materials}$

$$= \frac{N_{materials,space}}{N_{materials,QZP}} \cdot N_{material\ clusters,QZP} \cdot N_{\frac{observed}{cluster}} \cdot N_{\frac{total}{observed}} \cdot f_{use}$$

# What's left of quaternary space?



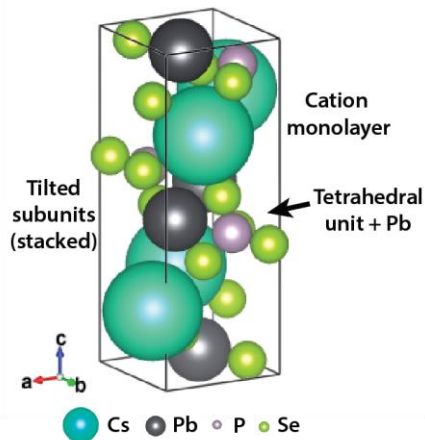
Sample a few interesting families.

# Families

Template :  $\text{CsPbPSe}_4$

Spacegroup:  $\text{Pnma}$

Stoichiometry: 1:1:1:4

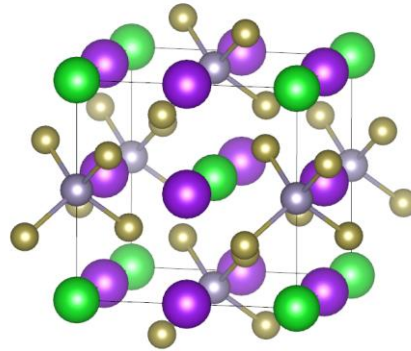


6 known

Template :  $\text{K}_2\text{BaSnTe}_4$

Spacegroup:  $\text{I-42m}$

Stoichiometry: 2:1:1:4

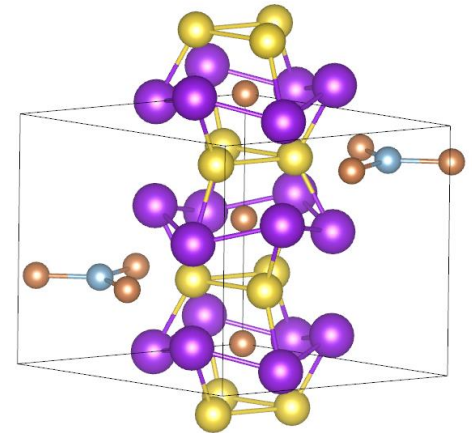


1 known

Template :  $\text{K}_6\text{AlSb}_4\text{Na}_3$

Spacegroup:  $\text{P63mmc}$

Stoichiometry: 6:1:4:3



4 known

# Creating new materials

A database of hypothetical materials is created

hydrogen 1 <b>H</b> 1.0079																	helium 2 <b>He</b> 4.0026						
lithium 3 <b>Li</b> 6.941	beryllium 4 <b>Be</b> 9.0122																	boron 5 <b>B</b> 10.811	carbon 6 <b>C</b> 12.011	nitrogen 7 <b>N</b> 14.007	oxygen 8 <b>O</b> 15.999	fluorine 9 <b>F</b> 18.998	neon 10 <b>Ne</b> 20.180
sodium 11 <b>Na</b> 22.990	magnesium 12 <b>Mg</b> 24.305																	aluminum 13 <b>Al</b> 26.982	silicon 14 <b>Si</b> 28.086	phosphorus 15 <b>P</b> 30.974	sulfur 16 <b>S</b> 32.065	chlorine 17 <b>Cl</b> 35.453	argon 18 <b>Ar</b> 39.948
potassium 19 <b>K</b> 39.098	calcium 20 <b>Ca</b> 40.078	scandium 21 <b>Sc</b> 44.956	titanium 22 <b>Ti</b> 47.867	vanadium 23 <b>V</b> 50.942	chromium 24 <b>Cr</b> 51.996	manganese 25 <b>Mn</b> 54.938	iron 26 <b>Fe</b> 55.845	cobalt 27 <b>Co</b> 58.933	nickel 28 <b>Ni</b> 58.693	copper 29 <b>Cu</b> 63.546	zinc 30 <b>Zn</b> 65.38	gallium 31 <b>Ga</b> 69.723	germanium 32 <b>Ge</b> 72.64	arsenic 33 <b>As</b> 74.922	selenium 34 <b>Se</b> 78.96	bromine 35 <b>Br</b> 79.904	krypton 36 <b>Kr</b> 83.798						
rubidium 37 <b>Rb</b> 85.468	strontium 38 <b>Sr</b> 87.62	yttrium 39 <b>Y</b> 88.906	zirconium 40 <b>Zr</b> 91.224	niobium 41 <b>Nb</b> 92.906	molybdenum 42 <b>Mo</b> 95.96	technetium 43 <b>Tc</b> [98]	ruthenium 44 <b>Ru</b> 101.07	rhodium 45 <b>Rh</b> 102.91	palladium 46 <b>Pd</b> 106.42	silver 47 <b>Ag</b> 107.87	cadmium 48 <b>Cd</b> 112.41	indium 49 <b>In</b> 114.82	tin 50 <b>Sn</b> 118.71	antimony 51 <b>Sb</b> 121.76	tellurium 52 <b>Te</b> 127.60	iodine 53 <b>I</b> 126.90	xenon 54 <b>Xe</b> 131.29						
caesium 55 <b>Cs</b> 132.91	barium 56 <b>Ba</b> 137.33	lanthanum 57 <b>La</b> 138.91	hafnium 72 <b>Hf</b> 178.49	tantalum 73 <b>Ta</b> 180.95	tungsten 74 <b>W</b> 183.84	rhenium 75 <b>Re</b> 186.21	osmium 76 <b>Os</b> 190.23	iridium 77 <b>Ir</b> 192.22	platinum 78 <b>Pt</b> 195.08	gold 79 <b>Au</b> 196.97	mercury 80 <b>Hg</b> 200.59	thallium 81 <b>Tl</b> 204.38	lead 82 <b>Pb</b> 207.2	bismuth 83 <b>Bi</b> 208.98	polonium 84 <b>Po</b> [209]	astatine 85 <b>At</b> [210]	radon 86 <b>Rn</b> [222]						



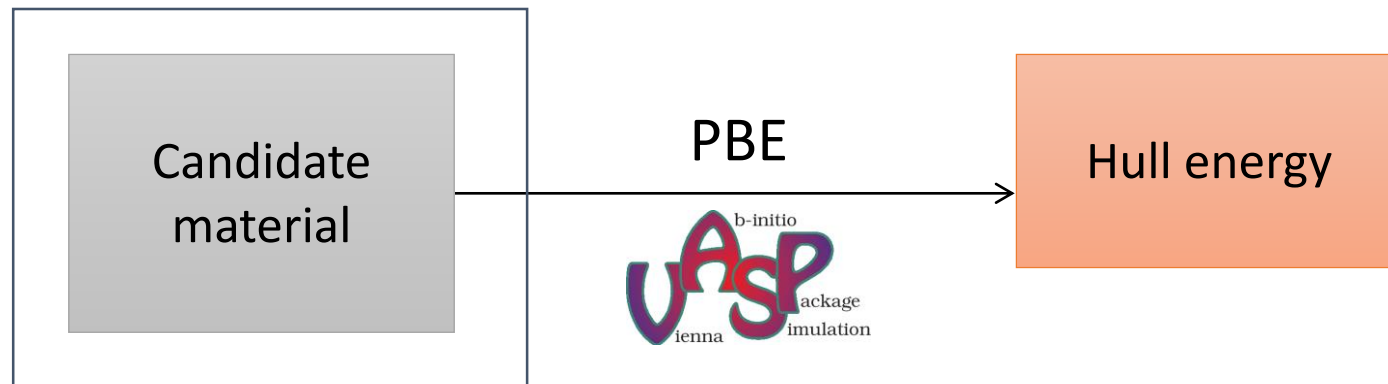
Combinations: 3168

2016

1696

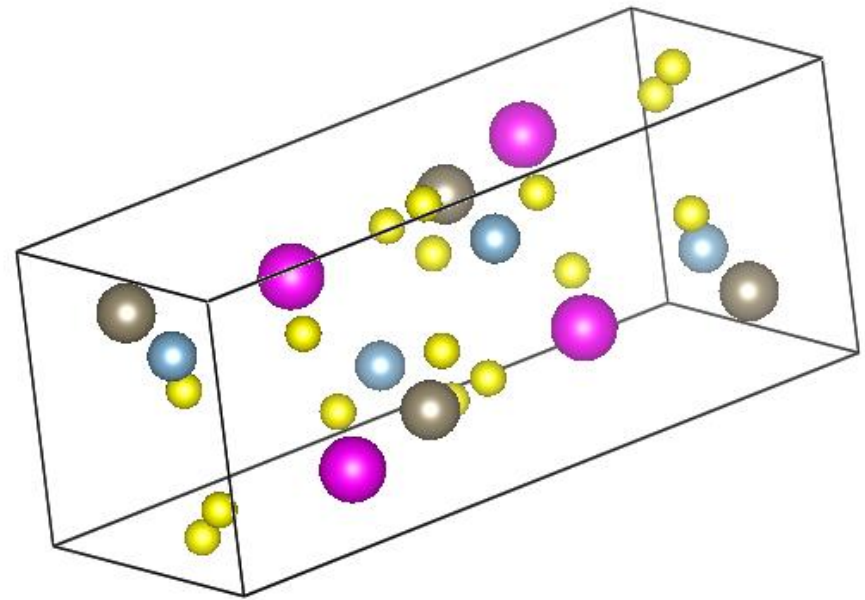
# Workflow

## High-throughput ab initio screening



# Workflow

BUILD STRUCTURE  
FROM  $\text{CsPbPSe}_4$

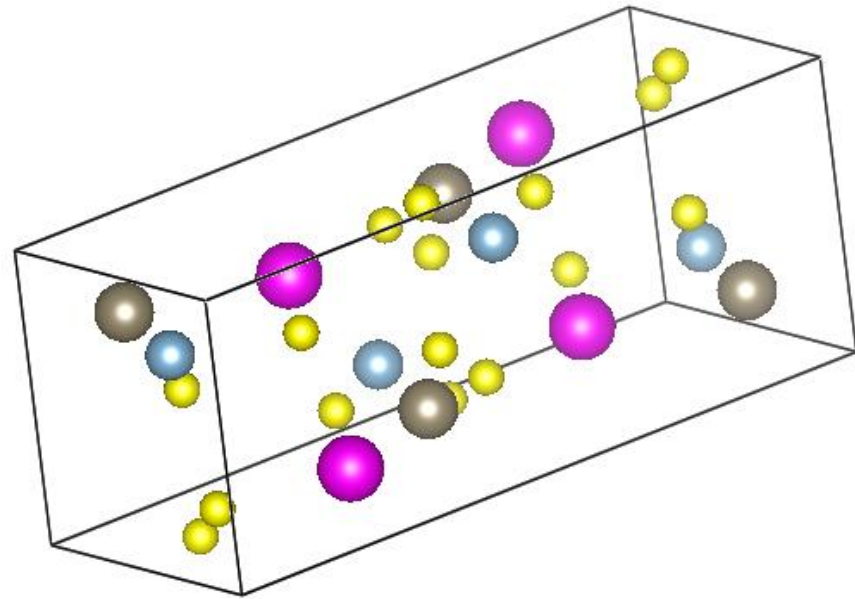




# Workflow

BUILD STRUCTURE  
FROM  $\text{CsPbPSe}_4$

VOLUME  
RELAXATION

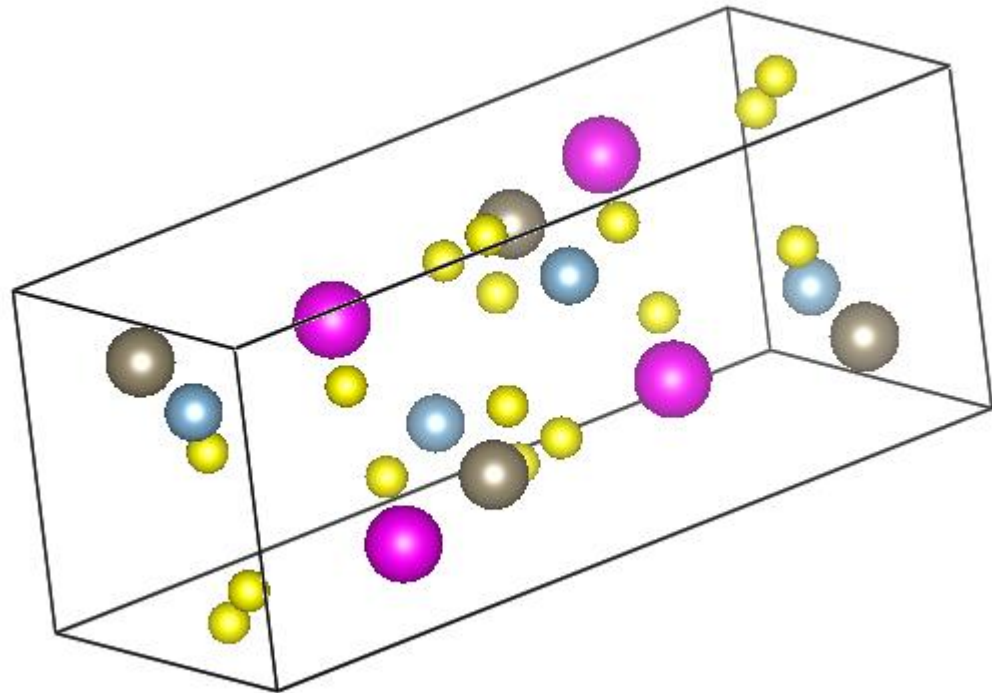


# Workflow

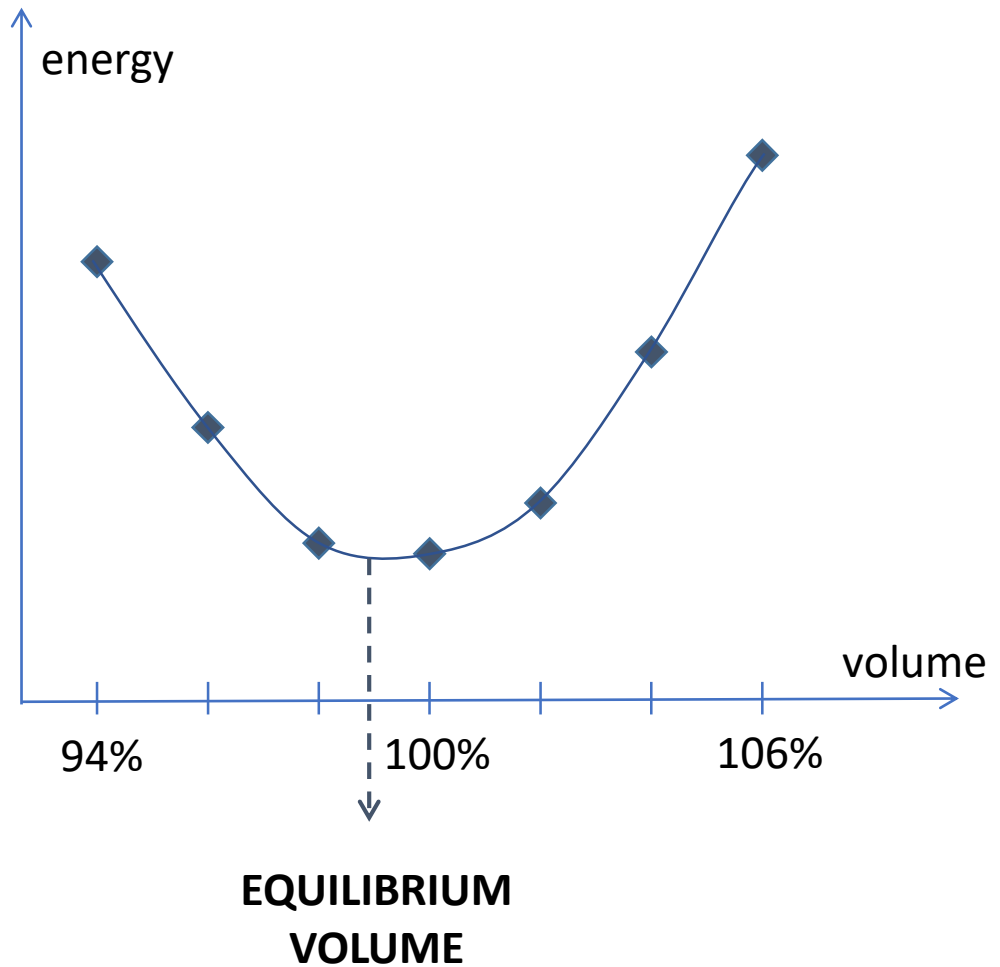
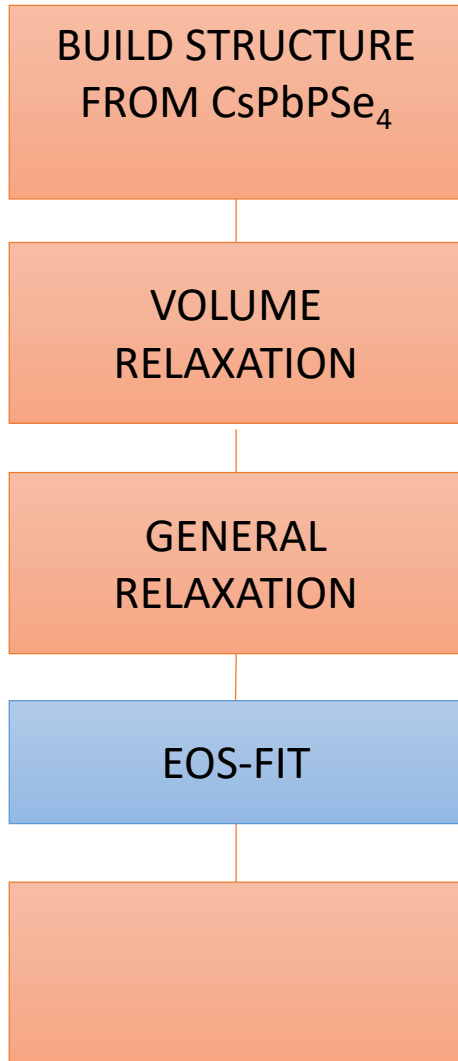
BUILD STRUCTURE  
FROM  $\text{CsPbPSe}_4$

VOLUME  
RELAXATION

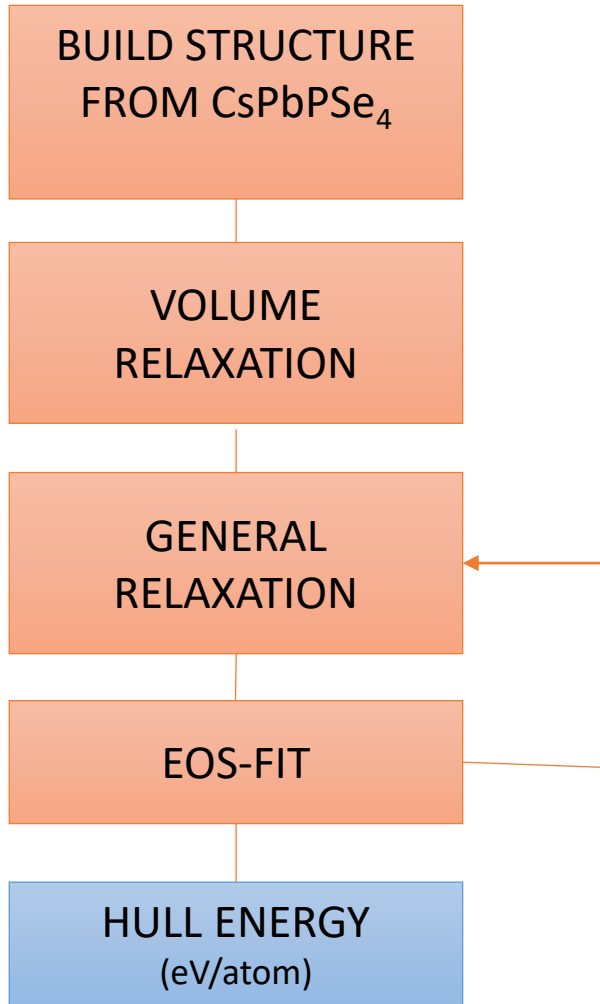
GENERAL  
RELAXATION



# Workflow



# Workflow



**+ - 20 calculations per material**



**> 100,000 calculations**

# Lost information

Most information is discarded

- Unstable materials: everything
- Stable materials: everything but the Ehull



**Recycle with machine learning!**



# Surrogate modeling

**Goal: learn the stable compositions as we perform calculations**

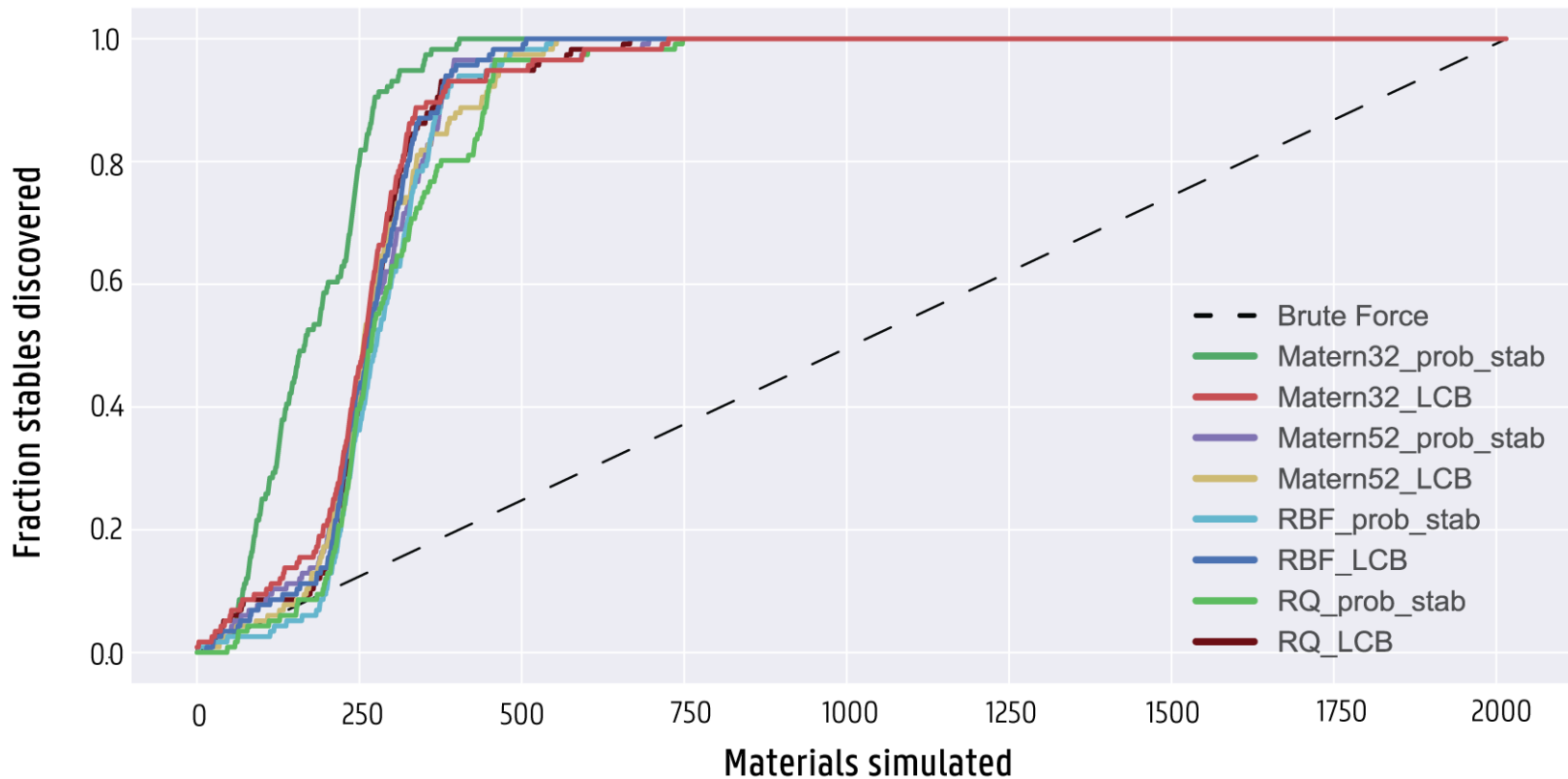


**Integrate with Queue Manager to prioritize calculations**

# Best model

Source: PhD work M. Larmuseau

## Best performance for a Matern32 kernel with probability sampling



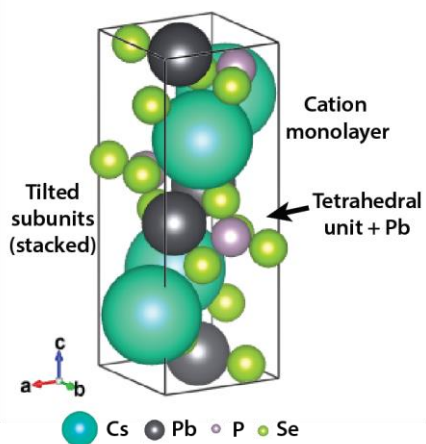
**90% of stable materials are found in 300 samples**

# How many did we find?

Template :  $\text{CsPbPSe}_4$

Spacegroup: Pnma

Stoichiometry: 1:1:1:4



108 stables

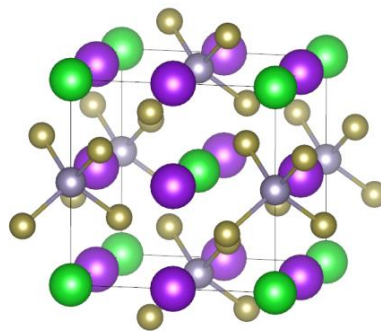
3168 total

1 in 30

Template :  $\text{K}_2\text{BaSnTe}_4$

Spacegroup: I-42m

Stoichiometry: 2:1:1:4



116 stables

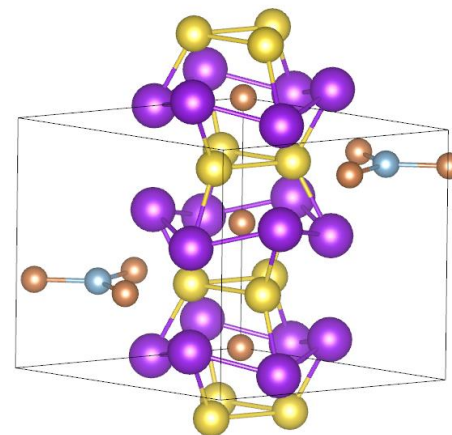
2016 total

1 in 20

Template :  $\text{K}_6\text{AlSb}_4\text{Na}_3$

Spacegroup: P63mmc

Stoichiometry: 6:1:4:3



115 stables

786 total

1 in 8



# Drake for quaternaries

## Materials Drake equation

$$N_{\text{useful materials}} = \frac{N_{\text{materials,space}}}{N_{\text{materials,QZP}}} \cdot N_{\text{clusters,QZP}} \cdot N_{\frac{\text{observed}}{\text{cluster}}} \cdot N_{\frac{\text{total}}{\text{observed}}} \cdot f_{\text{use}}$$

$$\frac{N_{\text{materials,space}}}{N_{\text{materials,QZP}}} = \frac{46399}{509} = 91.15$$

$$N_{\text{clusters,QZP}} = \pm 25 \text{ clusters} + 75 \text{ unique} \\ (+\text{partial occupancies})$$

$$N_{\frac{\text{observed}}{\text{cluster}}} N_{\frac{\text{total}}{\text{observed}}} = \pm 110$$

$$f_{\text{use}} = 1 \text{ (stable)}$$

The result is around 1 million with 47,000 known and 500 new ones found each year.

# Conclusions

- **High-throughput screening allows us to explore large material spaces**
- **Machine learning enables the recycling of information**
- **We can now find more stables in a month than experiment in a year**
- **Experimental synthesis is still needed**
- **Geometric models will allow us to reuse information not only within a family, but also between**