Variable and materials description

The database contains several properties of materials, that we will describe here. However, the first item of importance is that we considered for "our" database (the selection of a larger database) materials and properties answering to a few criteria:

- Only semi-conductors are considered. This refers to materials displaying a band of forbidden energy (a bandgap) for electric carriers. This is represented in the database under the form of the value of the bandgap (must be non-zero) when present.
- We restrict our study to materials with 2, 3 or 4 different atoms in their molecular formula. This is represented in the database by the molecular formula.
- Every entry in the database corresponds to a distinct compound, represented by its molecular formula. For this material, the property available in the database are then listed.
- **Each property entry in the database corresponds to an experimental measurement described in a peer reviewed article**. All property entries are associated with a link to the article detailing the experiment. However, since the access to the database is restricted, the link might simply appear blocked without due authorizations.
- When possible, the temperature at which was taken the measurement is also present. Since some properties are dependent on temperature, we tried to restrict the range of acceptable temperatures to 200-350K (-70°c to 80°c roughly).

Let's discuss now the different properties considered in building our sub-database.

- Enthalpy of formation (kJ g-at.$^{-1}$): change in enthalpy during the formation of a compound from its constituent elements. The enthalpy is a thermodynamical concept, defined as the internal energy of a system plus the volume multiplied by the pressure (H = U + P*V). It is always negative, and can be loosely interpreted as "how much does a material wants to form": the larger the absolute value of the enthalpy, the more likely a material is thermodynamically susceptible to form.
- Band gap (eV): in semi-conductors and insulators, the smallest energy necessary for an electron (resp. a hole) to make a transition from the valence band to the conduction band of a material (resp. conduction band to valence band). Also associated with light absorption, although in our case we mainly want to make sure that the materials studied have indeed a bandgap (if materials do not have a bandgap, E_G=0, then they are metals and do not enter the scope of the study).
- Heat capacity at constant pressure (J K^{-1} g-at.$^{-1}$): how much energy a material can store under the form of heat per unit of mass and temperature. This the first of the thermal properties investigated in the study.
- Thermal conductivity (W m^{-1} K^{-1}): ability of a material to conduct heat. Power transmitted in the form of heat per unit of length and temperature. Is linked to the propagation of vibration in the crystal in the form of phonons (quasi particles of lattice vibration). This is one of the properties where we could identify a clear distinction between the halides (materials containing at least an atom of Cl, Br or I) and other family of materials (see Fig 1).

- Linear thermal expansion coefficient (K^{-1}): how much a material expands (or contracts) when the temperature changes. The second thermal property where we could identify a clear separation of the families of materials (see Fig 1).

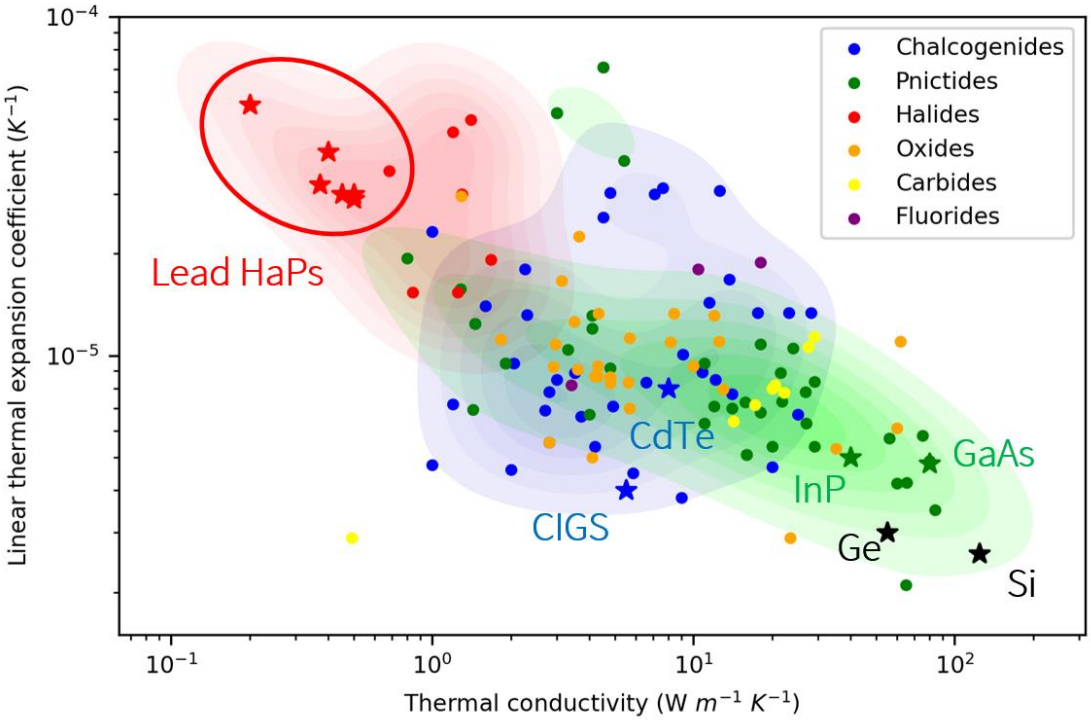

Figure 1: Thermal conductvity vs linear thermal expansion coefficient, data from the MPDS database under the restrictions detailed above

The thermal conductivity vs linear thermal expansion coefficient figure tells us that halide (and lead halide perovskite, HaPs, in particular) are materials that are pretty bad at conducting heat. This is a known phenomenon for the lead HaPs, but here it is clearly confirmed on a "large" number of semi-conductors. The high thermal expansion coefficient on the other hand tells us that halide and lead HaPs in particular are materials that are subjects to large changes in terms of crystal structure when the temperature changes. For a deeper analysis of this existing relationship and see how these measurements are taken, see the article from Ge et al (DOI: 10.1021/acs.jpcc.8b05919). This paints a picture of soft materials, easily deformed. This will be confirmed by our next set of properties.

Let's move to what could be considered more mechanical properties of materials:

- Sound velocity (m s^{-1}): the speed of sound in the material. Linked to the density of the material (also accessible in the database). No obvious correlation observed between sound velocity and composition.
- Microhardness (GPa): the ability of a material to resist to deformations under constraints, at a microscopical level. A high microhardness means that the material is difficult to deform, in other words hard. This is actually considered a "macro" property (large scale property, in opposition with "micro" properties that can usually be defined at the scale of a few unit cells of the crystal

lattice). One of the two mechanical properties where we could once again identify a clear distinction between families of materials (see Fig 2).

- Isothermal bulk modulus (GPa): the capacity of the crystal lattice of a material to resist to deformation under pressure at constant temperature. The second mechanical property where we could identify a distinction between families of materials (see Fig 2).

In the case of mechanical properties, halides (and lead halide perovskites in particular) seem to be particularly soft materials (low bulk modulus and microhardness).

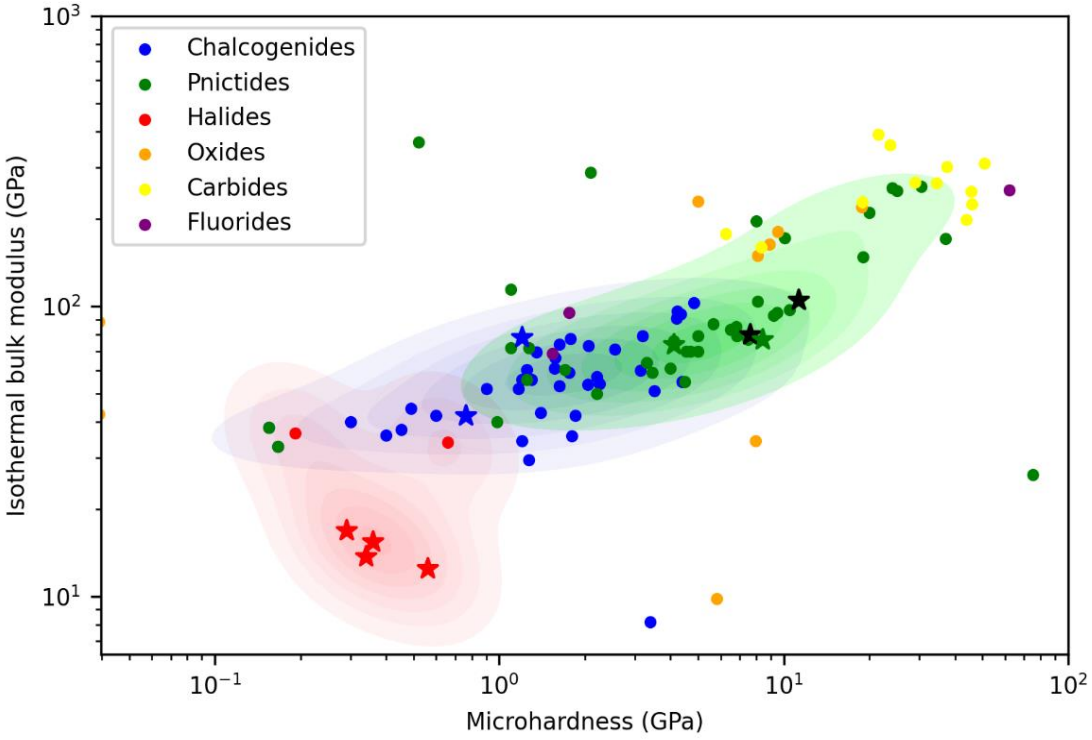

Figure 2: Microhardness vs Isothermal bulk modulus, data from the MPDS database under the restrictions detailed above. Red stars represent lead halide perovskites.