Atomis'c Materials Design Using the Computational Materials Repository

Karsten W. Jacobsen
Technical University of Denmark

Follow this and additional works at: http://dc.engconfintl.org/materials_genome

Part of the Biomedical Engineering and Bioengineering Commons

Recommended Citation

This Conference Proceeding is brought to you for free and open access by the Proceedings at ECI Digital Archives. It has been accepted for inclusion in Harnessing The Materials Genome: Accelerated Materials Development via Computational and Experimental Tools by an authorized administrator of ECI Digital Archives. For more information, please contact franco@bepress.com.
Atomistic Materials Design Using the Computational Materials Repository

Karsten W. Jacobsen

Center for Atomic-scale Materials Design
Dept. of Physics
Technical University of Denmark
Outline

• Computational Materials Repository
• Perovskites for water splitting
• New optimized exchange-correlation density functionals
Computational Materials Repository

- A system for storing/uploading, analyzing, retrieving, and sharing computational data.
  - Some ideas:
    - Many interfaces (sql, python, web, ”silo”)
    - Agents – small pieces of code automatically performing calculations in the database
    - Taxonomy/folksonomy
    - Data identification – publication
  - Software at [http://wiki.fysik.dtu.dk/cmr](http://wiki.fysik.dtu.dk/cmr)
  - Data at [https://cmr.fysik.dtu.dk](https://cmr.fysik.dtu.dk)
Generic data view

<table>
<thead>
<tr>
<th>id_ref</th>
<th>jmol</th>
<th>A</th>
<th>B</th>
<th>anion</th>
<th>heat_of_formation</th>
<th>gilbse_dir_gap</th>
<th>gilbse_ind_gap</th>
<th>Cil_dir</th>
<th>Cil_ind</th>
<th>VD_dir</th>
<th>VD_ind</th>
<th>atoms</th>
<th>db_keywords</th>
<th>doi</th>
<th>ingredients</th>
<th>downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>8926</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0400</td>
<td>0</td>
<td>0</td>
<td>5.9047</td>
<td>5.9247</td>
<td>5.9747</td>
<td>5.9747</td>
<td>0.10</td>
<td>Re Rh Re (30, 1Rh)</td>
<td>10.1039/C1EE027170</td>
<td>Ag03 Ag max O 0.25 RM + 3.00 O</td>
<td>177</td>
</tr>
<tr>
<td>8935</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.3700</td>
<td>0</td>
<td>0</td>
<td>6.0569</td>
<td>6.0569</td>
<td>6.0569</td>
<td>6.0569</td>
<td>0.10</td>
<td>Ag Ag (30, 1Rh)</td>
<td>10.1039/C1EE027170</td>
<td>Ag03 Ag max O 0.25 RM + 3.00 O</td>
<td>172</td>
</tr>
<tr>
<td>8938</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5200</td>
<td>0</td>
<td>0</td>
<td>6.0182</td>
<td>6.0182</td>
<td>6.0182</td>
<td>6.0182</td>
<td>0.10</td>
<td>Cu Rh (30, 1Cu, 1Rh)</td>
<td>10.1039/C1EE027170</td>
<td>Ag03 Cu max O 0.25 RM + 3.00 O</td>
<td>175</td>
</tr>
<tr>
<td>8945</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.9100</td>
<td>0</td>
<td>0</td>
<td>6.0677</td>
<td>6.0677</td>
<td>6.0677</td>
<td>6.0677</td>
<td>0.10</td>
<td>Cu Cu (30, 1Cu, 1Cu)</td>
<td>10.1039/C1EE027170</td>
<td>Ag03 Cu max O 0.25 RM + 3.00 O</td>
<td>177</td>
</tr>
<tr>
<td>8946</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.1400</td>
<td>0</td>
<td>0</td>
<td>6.1417</td>
<td>6.1417</td>
<td>6.1417</td>
<td>6.1417</td>
<td>0.10</td>
<td>Si Rh (30, 1Si, 1Rh)</td>
<td>10.1039/C1EE027170</td>
<td>Ag03 Si max O 0.25 RM + 3.00 O</td>
<td>177</td>
</tr>
</tbody>
</table>
Project specific interface:
Light absorbing materials for water splitting

http://cmr.fysik.dtu.dk - the database
Water splitting

Complicated process:
- Light absorption
- Electron-hole motion
- Induce reactions

Examples: TiO$_2$, GaN:ZnO, ZnGeN$_2$:ZnO

(Fujishima and Honda, Nature 1972)
(Maeda et al., JACS, 127, 8286 (2005))
Materials for water splitting

- Chemical/structural stability
- Band gap of 1.5-3 eV (overpotentials, losses)
- Band edge positions straddle the water redox potentials
- Good electron/hole mobilities
- Low cost, non-toxicity
- Good catalytic properties
Methodology – density functional theory + friends

GPAW – projector augmented wave method in real space

https://wiki.fysik.dtu.dk/gpaw/ ← Free download, GPL

- High accuracy: Wave functions expanded on real space grids or plane waves
- High efficiency: Wave functions expanded in atomic-like orbitals (LCAO)
- Efficient parallelization (good scalability up to > 32,000 CPUs)
- Xc-functionals: LDA, GGAs, meta-GGA, LDA+U, EXX, vdWDF, GLLB, BEEF
- Time-dependent DFT (including “Bootstrap”)
- Many-body perturbation theory (GW and Bethe-Salpeter equation)
- Phonons and electron-phonon coupling
- Quantum electron transport
- Atomic Simulation Environment (ASE) python scripting interface
Materials – cubic perovskites

• Perovskite, common stable structure, 50% are quasi-cubic
• Variety of properties: ferroelectricity, magnetism, superconductivity and (photo)catalytic activity
• 52 different metallic elements
• Different anions (O, N, S, F, Cl, …)

Excluded elements:
• Non Metals;
• Radioactive, toxic.
Predicting stability of oxides – Heat of formation

- Focus on oxides because of high stability (towards oxidation!)
- DFT-RPBE calculated formation energy for rutile dioxides.
- Similar results obtained for perovskite structures.

\[ \Delta G_{\text{Form}} \]

\[ MAE = \frac{1}{n} \sum_{i} |y_i - x_i| = 0.29 \text{ eV} \]
Calculation of bandgaps

- DFT is aimed at calculating ground state total energies – does not provide bandgap
- Bandgaps particularly bad for (semi-)local approximations
- GLLB approximation
  - Improved xc-potential compared to LDA/GGA
  - Explicit evaluation of derivative discontinuity
    \[ E_{g}^{QP} = E_{g}^{KS} + \Delta_{xc} \]

Optical absorption spectra with GLLB-SC

Derivative discontinuity used in spectrum for TDDFT, but not for W in BSE.

Ag surface plasmon with GLLB-SC

J. Yan, K. W. Jacobsen, and K. S. Thygesen, PRB 84, 235430 (2011)
The GLLB-SC (solid-correlation) xc-functional:

\[ E_g^{QP} = E_g^{KS} + \Delta_{xc} \]

Derivative discontinuity

- Bandgaps within \( \sim 0.5 \) eV of exp.
- Minimal computational cost
- Neglect of electron-hole interaction – excitons
\[ E_C = \left( \chi_A \chi_B \chi_O^3 \right)^{1/5} - \frac{1}{2} E_{\text{gap}} + E_0 \]

\[ \chi = \frac{1}{2}(A+I_1) \]

Absolute electronegativity (Mulliken scale)

A = Affinity level

I_1 = Ionization level

\( E_{\text{gap}} = \) Band gap

\( E_0 = \) Difference between NHE and vacuum \( \sim -4.5 \) eV

Empirical formula: 

\[ \frac{1}{2} O_2 + 2 H^+ \rightarrow H_2 + O_2 \]

\( H^+/H_2 \) Potential

\( O_2/H_2O \) Potential


Y Xu and MAA Schoonen, American Mineralogist (2000)
One- and two-photon water splitting

### One-photon

- **H₂**
- **H₂O**

**O₂** and **H₂** evolution photocatalyst

**O₂/H₂O** Potential

**H⁺/H₂** Potential

### Two-photon

- **H₂**
- **2 H⁺**

**½ O₂ + 2 H⁺**

**H₂** evolution photocatalyst

**O₂** evolution photocatalyst

**O₂/H₂O** Potential

**H⁺/H₂** Potential

### Screening parameters

<table>
<thead>
<tr>
<th></th>
<th>One-photon WS</th>
<th>Two-photon WS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical/structural stability (ΔE)</td>
<td>ΔE \leq 0.2 eV</td>
<td>ΔE \leq 0.2 eV</td>
</tr>
<tr>
<td>Bandgap (E_{gap})</td>
<td>1.5 \leq E_{gap} \leq 3 eV</td>
<td>1.3 \leq E_{gap} \leq 3 eV</td>
</tr>
<tr>
<td>Band edges (VB_{edge}, CB_{edge})</td>
<td>VB_{edge} &gt; 1.23 eV, CB_{edge} &lt; 0 eV</td>
<td>VB_{anode}^{cathode} &gt; CB_{edge}^{anode}</td>
</tr>
</tbody>
</table>

H₂ photocatalyst: Si

O₂ photocatalyst: screening
Tandem cell efficiency

Solar-to-hydrogen energy conversion efficiency

(J. R. Bolton et al., Nature 1985.)
(M. G. Walter et al., Chem Rev 110, 6446, 2010)
Perovskites: Stability vs. band gap

+ many highly stable metallic systems

~19000 materials

Bandgap distribution for ABO$_3$
Perovskites: ABO$_3$ candidates

**Stability:** formation energy < 0.2 eV;

**Light absorption:** 1.5 eV < band gap < 3 eV.

43 oxides

↓

Extended stability

13 oxides

↓

Level alignment

10 oxides

**Heat of Formation (eV)**

**Energy Gap (eV)**
One-photon water splitting – oxide candidates

Empirical formula for the conduction band relative to NHE:

$$E_C = (\chi_A \chi_B \chi_O)^{3/5} - 1/2 E_{gap} + E_0$$

Butler and Ginley (1978)

AgNbO$_3$ and BaSnO$_3$ known.

AgNbO$_3$ works!

BaSnO$_3$ defect-induced recombination

SrSnO$_3$ and CaSnO$_3$:

known in orthorhombic perovskite

→ too large gaps

10 materials identified
Oxides, oxynitrides, oxysulfides, oxyfluorides, oxyfluorinitrides

Materials candidates:

- **ABO$_3$** : 10
  - BaTaO$_3$N, SrTaO$_3$N, CaTaO$_2$N, LaTiO$_2$N (known)
  - MgTaO$_2$N (unknown)
- **ABO$_2$N** : 5
  - LaTaON$_2$ (known)
  - YTaON$_2$ (unknown)
- **ABON$_2$** : 2
- **ABN$_3$** : 0
- **ABO$_2$S** : 0
- **ABO$_2$F** : 3
- **ABOFN** : 0

$\sim$19000 materials

One-photon water splitting

20 candidate materials
Further analysis of candidate materials: bandgap calculations

In the future:
Absorption spectra

PBE0: hybrid
GW: $G_0W_0$ + plasmon pole
Tandem cell water splitting: Screening results

12 candidates
+ 20 from overall WS

E_{form} < 0.2
1.3 < E_{gap} < 3 \text{ eV}

LaTiO$_2$N now under experimental investigation at CINF/ CASE/DTU.

Data mining

• Understanding
  • Wealth of information available
  • Many (and identifying new) chemical concepts and rules unused
    • Valence
    • Metal, if odd number of electrons in unit cell
    • Stability size rule: tolerance factor \sim 1

• Prediction
  • Use information from one screening to make subsequent ones more efficient
    • Maybe based on understanding
    • Maybe not – “machine learning”
Clusters follow the valences of the elements.

46 stable $\text{ABO}_3$ showing bandgaps

Clusters based on bandgap for $\text{ABO}_3$: rediscovering valence

Zr, Hf group IV

Ca, Sr, Ba (valence 2)

Ga, In Tl (group III)

Alkalis + Ag (valence 1)
Probability for a perovskite with a given A-ion (or B-ion) to be stable and have a bandgap considering only systems obeying the two rules:

Odd/even rule: only a system with even number of electrons in the unit cell can form a semiconductor/insulator.

Valence rule: the sum of the possible valences of the two metals and of the anions have to be equal in absolute value for a semiconductor to be possible.

### ABO$_3$

<table>
<thead>
<tr>
<th>A-ion</th>
<th>Probability [%]</th>
<th>B-ion</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na</td>
<td>60.0</td>
<td>Ta</td>
<td>34.6</td>
</tr>
<tr>
<td>Sr</td>
<td>50.0</td>
<td>Nb</td>
<td>34.6</td>
</tr>
<tr>
<td>Ca</td>
<td>50.0</td>
<td>Zr</td>
<td>33.3</td>
</tr>
<tr>
<td>Li</td>
<td>50.0</td>
<td>Hf</td>
<td>27.8</td>
</tr>
<tr>
<td>K</td>
<td>50.0</td>
<td>Ti</td>
<td>26.1</td>
</tr>
<tr>
<td>Ba</td>
<td>40.0</td>
<td>Sn</td>
<td>13.6</td>
</tr>
<tr>
<td>Cs</td>
<td>33.3</td>
<td>Al</td>
<td>10.5</td>
</tr>
<tr>
<td>Rb</td>
<td>33.3</td>
<td>Ge</td>
<td>9.1</td>
</tr>
<tr>
<td>Ag</td>
<td>33.3</td>
<td>Sb</td>
<td>7.7</td>
</tr>
<tr>
<td>La</td>
<td>21.1</td>
<td>V</td>
<td>7.4</td>
</tr>
<tr>
<td>Sn</td>
<td>13.6</td>
<td>Sc</td>
<td>5.2</td>
</tr>
<tr>
<td>Pb</td>
<td>13.6</td>
<td>Ga</td>
<td>5.2</td>
</tr>
<tr>
<td>Tl</td>
<td>10.5</td>
<td>In</td>
<td>5.2</td>
</tr>
<tr>
<td>In</td>
<td>10.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ga</td>
<td>10.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td>10.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ge</td>
<td>4.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### ABO$_2$N

<table>
<thead>
<tr>
<th>A-ion</th>
<th>B-ion</th>
<th>Stable?</th>
<th>Bandgap [eV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ca</td>
<td>Ta</td>
<td>✓</td>
<td>2.2</td>
</tr>
<tr>
<td>Sr</td>
<td>Ta</td>
<td>✓</td>
<td>2.1</td>
</tr>
<tr>
<td>Ca</td>
<td>Nb</td>
<td>✓</td>
<td>1.4</td>
</tr>
<tr>
<td>Sr</td>
<td>Nb</td>
<td>✓</td>
<td>1.4</td>
</tr>
<tr>
<td>Ba</td>
<td>Ta</td>
<td>✓</td>
<td>2.0</td>
</tr>
<tr>
<td>La</td>
<td>Ti</td>
<td>✓</td>
<td>2.5</td>
</tr>
<tr>
<td>Ba</td>
<td>Nb</td>
<td>✓</td>
<td>1.3</td>
</tr>
<tr>
<td>Pb</td>
<td>Ta</td>
<td>✓</td>
<td>2.0</td>
</tr>
<tr>
<td>La</td>
<td>Zr</td>
<td>✓</td>
<td>3.4</td>
</tr>
<tr>
<td>Pb</td>
<td>Nb</td>
<td>✓</td>
<td>1.3</td>
</tr>
<tr>
<td>La</td>
<td>Hf</td>
<td>✓</td>
<td>3.8</td>
</tr>
<tr>
<td>Mg</td>
<td>Ta</td>
<td>✓</td>
<td>2.1</td>
</tr>
<tr>
<td>Mg</td>
<td>Nb</td>
<td>✓</td>
<td>1.5</td>
</tr>
<tr>
<td>Ca</td>
<td>V</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Sr</td>
<td>V</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>La</td>
<td>V</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>La</td>
<td>Sn</td>
<td></td>
<td>1.8</td>
</tr>
<tr>
<td>Ga</td>
<td>Ti</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Y</td>
<td>Ti</td>
<td></td>
<td>2.4</td>
</tr>
<tr>
<td>In</td>
<td>Ti</td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Tl</td>
<td>Ti</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Ba</td>
<td>V</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>La</td>
<td>Si</td>
<td></td>
<td>3.0</td>
</tr>
<tr>
<td>La</td>
<td>Ge</td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Ge</td>
<td>Ta</td>
<td></td>
<td>1.8</td>
</tr>
<tr>
<td>Sn</td>
<td>Ta</td>
<td>✓</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Towards bandgap engineering: double perovskites

Double perovskite: obtained combining two stable cubic perovskites with a bandgap.

New “design rules”:
• Double perovskite has average of perovskite gaps
• But, for B1 p-metal and B2 d-metal gap is significantly increased (but typically > 4 eV)
• Several hundred new potential water splitting materials discovered
Towards bandgap engineering: layered perovskites + ICSD

Preliminary screenings:

Collaboration with Materials Project, Anubhav Jain, Gerbrand Ceder, Bandgaps calculated for pre-optimized structures
CMR website for perovskites

**Computational Materials Repository**

Do not forget to press update matrix after changing the selection! If there is an error - it means that the dataset is already being calculated! Please wait a moments and try again.

- Chose a data set: ABO3 (2704)
- Width: 800
- Height: 1200
- X axis ticks: automatically selected
- Y axis ticks: automatically selected
- X sort order: Electronegativity (Pauling)
- Y sort order: Electronegativity (Pauling)
- Action on Click: show band edges
- References:
  - ABN (3)
  - ABO (20)
  - AN (50)
  - AO (52)
  - AON (35)
  - default (3)
  - mbulk (52)

**Value field: Colors:**

- Triangle 1: (top-right) gilbsc_ind-gap (eV)
- Triangle 2: (bottom-left) heat_of_formation (eV)
- Triangle 3:
- Triangle 4:

**Examples for the color choice:**

- 0=>white,1=>red,7=>blue
- -100=>blue,100=>red

Valid color names are black, blue, cyan, green, gray, green, lightblue, pink, red, purple, white, yellow. Please note that the values must be in increasing order.

**TITaO3**

Heat of Form. = 0.1 eV/atom

- Indirect Gap = 2.0 eV
- Direct Gap = 2.0 eV

- Valence Band = 2.8 (2.8) eV
- Center Band = 1.8 eV
- Conduction Band = 0.8 (0.8) eV

---

[http://cmr.fysik.dtu.dk](http://cmr.fysik.dtu.dk) - the database

Optimized Bayesian Error Estimation xc-Functional: BEEF-vdW

\[ E_{xc} = \sum_{m=0}^{\infty} a_m E_m^{\text{GGA-x}} + \alpha_c E^{\text{LDA-c}} + (1 - \alpha_c) E^{\text{PBE-c}} + E^{\text{nl-c}} \]

- Linear model
  - GGA exchange expanded on orthogonal Legendre polynomials
  - Non-local correlation included as in vdW-DF2

- Databases
  - Molecules (fragmentation, reaction energies and barriers)
  - Solids (cohesive energies, lattice constants)
  - Chemisorption energies
  - Binding energies for non-covalently bonded systems

- Important issues
  - Avoid overfitting – Tikhonov regularization, bootstrapping
  - Balancing relative importance of datasets

BEEF-vdW:
Here is the BEEF(-vdW)

Available in the GPAW code including error estimation
Overall very good performance. “Robust”.

Mean Absolute Deviation

G3: molecular formation energies
CE27: chemisorption on TM surfaces
DBH24: gas-phase reaction barriers
RE42: gas-phase reaction energies
Sol27Ec: solid cohesive energies
Sol27Lc: solid lattice constants
S22x5: non-covalent (vdW) binding
Breaking the “GGA-line”?

CO/Rh(111) adsorption energy vs. the Rh(111) surface energy.

Error bars indicate BEEF-vdW ensemble error estimates.

Purple diamond indicates experimental result.

Coming up: Functional at meta-GGA level.

Original plot from:
Data available in CMR

COMPUTATIONAL MATERIALS REPOSITORY

Access to data from the study:

Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation

http://cmr.fysik.dtu.dk

<table>
<thead>
<tr>
<th>name</th>
<th>atoms</th>
<th>ref_energy</th>
<th>PBE</th>
<th>RPBE</th>
<th>PBEsol</th>
<th>revTPSS</th>
<th>vdW-DF</th>
<th>vDW-DF2</th>
<th>optB88-vdW</th>
<th>BEEF-vdW</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2/Pl(111)</td>
<td>H (1H, 20Pd)</td>
<td>-0.410</td>
<td>-0.976</td>
<td>-0.657</td>
<td>-1.348</td>
<td>-0.816</td>
<td>-0.330</td>
<td>-0.316</td>
<td>-0.777</td>
<td>-0.462</td>
</tr>
<tr>
<td>H2/Ni(111)</td>
<td>Ni (1H, 20Ni)</td>
<td>-0.980</td>
<td>-1.131</td>
<td>-0.786</td>
<td>-1.526</td>
<td>-1.108</td>
<td>-0.581</td>
<td>-0.530</td>
<td>-0.996</td>
<td>-0.687</td>
</tr>
<tr>
<td>H2/Ni(100)</td>
<td>Ni (1H, 20Ni)</td>
<td>-0.930</td>
<td>-1.053</td>
<td>-0.737</td>
<td>-1.519</td>
<td>-1.137</td>
<td>-0.439</td>
<td>-0.338</td>
<td>-0.915</td>
<td>-0.602</td>
</tr>
<tr>
<td>H2/Rh(111)</td>
<td>Rh (1H, 20Rh)</td>
<td>-0.810</td>
<td>-1.067</td>
<td>-0.725</td>
<td>-1.431</td>
<td>-0.986</td>
<td>-0.482</td>
<td>-0.423</td>
<td>-0.882</td>
<td>-0.618</td>
</tr>
<tr>
<td>H2/Pd(111)</td>
<td>Pd (1H, 20Pd)</td>
<td>-0.910</td>
<td>-1.117</td>
<td>-0.806</td>
<td>-1.503</td>
<td>-0.972</td>
<td>-0.491</td>
<td>-0.465</td>
<td>-0.941</td>
<td>-0.627</td>
</tr>
<tr>
<td>H2/Ir(111)</td>
<td>Ir (1H, 20Ir)</td>
<td>-0.550</td>
<td>-0.852</td>
<td>-0.535</td>
<td>-1.250</td>
<td>-0.836</td>
<td>-0.283</td>
<td>-0.247</td>
<td>-0.707</td>
<td>-0.426</td>
</tr>
<tr>
<td>H2/Co(O001)</td>
<td>Co H (20Co, 1H)</td>
<td>-0.690</td>
<td>-1.131</td>
<td>-0.813</td>
<td>-1.516</td>
<td>-1.128</td>
<td>-0.602</td>
<td>-0.569</td>
<td>-1.006</td>
<td>-0.699</td>
</tr>
<tr>
<td>H2/Ru(O001)</td>
<td>Ru H (1H, 20Ru)</td>
<td>-1.040</td>
<td>-1.197</td>
<td>-0.884</td>
<td>-1.555</td>
<td>-1.192</td>
<td>-0.610</td>
<td>-0.591</td>
<td>-1.014</td>
<td>-0.739</td>
</tr>
<tr>
<td>N2/Fe(100)</td>
<td>Fe N (20Fe, 2N)</td>
<td>-2.300</td>
<td>-2.885</td>
<td>-2.050</td>
<td>-3.953</td>
<td>-3.386</td>
<td>-2.138</td>
<td>-2.088</td>
<td>-3.190</td>
<td>-1.986</td>
</tr>
</tbody>
</table>

http://cmr.fysik.dtu.dk
Outlook

• Computational Materials Repository
  • MGI – much more software development needed
  • A GPL standard for software and data?
• Water splitting materials screening
  • Tailoring of bandgaps
  • Additional important factors
    • Carrier mobilities
    • Catalysis
• XC functional development
  • Important for efficient calculations
  • Much larger expt/comp databases needed – coupled cluster/RPA+more – more systematic development
  • More sophisticated machine learning techniques
Acknowledgements

CAMD/DTU:
Ivano E. Castelli
Thomas Olsen
Kristian S. Thygesen
Jun Yan
Soumendu Datta (now at Bose Ctr, Kalkutta)
Vivien Petzold
Keld Lundgaard
Jess Wellendorff
David D. Landis
Marcin Dulak

SLAC/Stanford Univ.:
Andreas Møgelhøj
Jens Hummelshøj
Thomas Bligaard
Jens K. Nørskov

CINF-CASE/DTU:
Ib Chorkendorff
Søren Dahl (now at Topsøe A/S)

Stanford Univ:
Tom Jaramillo

Argonne Nat’l Lab:
Jeff Greeley

Univ. of Chicago:
Svetlozar Nestorov
Computational Materials Repository

The database: http://cmr.fysik.dtu.dk


Optimized xc-functionals


Watersplitting


Bandgap calculations with GLLB

The GLLB xc-functional (Gritsenko, van Leeuwen, van Lenthe and Baerends):

\[ E_g^{QP} = E_g^{KS} + \Delta_{xc} \]

Derivative discontinuity

\[
v_x(r) = v_S(r) + v_{\text{resp}}(r)
\]

Screening + response

\[
v_S(r) = \frac{2\varepsilon_x^{\text{GGA}}(r; n)}{n(r)}
\]

\[
v_{\text{resp}}(r) = \sum_{\text{occ}} K[n] \sqrt{\varepsilon_r - \varepsilon_i} \frac{|\psi_i(r)|^2}{n(r)}
\]

\[
\Delta_{x,\text{resp}}(r) = \sum_i^N K (\sqrt{\varepsilon_{N+1}} - \varepsilon_i - \sqrt{\varepsilon_N} - \varepsilon_i) \frac{|\psi_i(r)|^2}{n(r)}
\]

GLLB-SC: Screening exchange-correlation from PBEsol

<table>
<thead>
<tr>
<th>Material</th>
<th>( E_g^{KS} ) (LDA)</th>
<th>( E_g^{KS} )</th>
<th>( \Delta_{xc} )</th>
<th>( E_g^{QP} )</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>4.09</td>
<td>4.14</td>
<td>1.27</td>
<td>5.41</td>
<td>5.48</td>
</tr>
<tr>
<td>Si</td>
<td>0.443</td>
<td>0.68</td>
<td>0.32</td>
<td>1.00</td>
<td>1.17</td>
</tr>
<tr>
<td>GaAs</td>
<td>0.36</td>
<td>0.79</td>
<td>0.25</td>
<td>1.04</td>
<td>1.63</td>
</tr>
<tr>
<td>AlAs</td>
<td>1.34</td>
<td>1.67</td>
<td>0.82</td>
<td>2.49</td>
<td>2.32</td>
</tr>
<tr>
<td>LiF</td>
<td>8.775</td>
<td>10.87</td>
<td>4.09</td>
<td>14.96</td>
<td>14.2</td>
</tr>
<tr>
<td>Ar</td>
<td>8.18</td>
<td>10.28</td>
<td>4.69</td>
<td>14.97</td>
<td>14.2</td>
</tr>
</tbody>
</table>


Perovskites: Heat of formation

Stable materials:
- Low electronegativity
- Sum of oxidation numbers = 6
- Geometric tolerance factor ~1
Most perovskites are metallic or low-gap semiconductors.
Analyzing gap formation

**ZnSiO₃**
- Formation energy = -1 eV;
- Band gap = 2.4 eV.
- Valence band: **O – p orbitals** (too deep for water-splitting);
- Conduction band: **Zn – s orbitals**.

**AgNbO₃**
- Formation energy = -0.6 eV;
- Band gap = 3.0 eV.
- Valence band: **Ag – d** and **O – p orbitals**;
- Conduction band: **Nb – d orbitals**.
Oxynitrides

**BaTaO_2N**

Formation energy = -6.3 eV; 
Band gap = 2.0 eV.

Valence band: 
\( Ta - p \) and \( N - p \) orbitals;

Conduction band: 
\( Ta - d \) orbitals.
Requirements:

- structural/chemical stability;
- two visible light harvests (optimal gaps: 1.1 eV and 1.7 eV);
- band edges that match with oxygen and hydrogen potentials;
- Small overlap between the semiconductors band edges for the electron transfer reaction.

Two semiconductors – two photons

- SC 1: Hole for oxygen evolution
- SC 2: Electron for hydrogen evolution

H₂ photocatalyst: Si
O₂ photocatalyst: screening
Pool of reference systems:

- Single metal bulk: A(s) and B(s)
- Single metal oxides: $A_xO_y(s)$ (and nitrides, sulfides, ...)
- Bimetallic oxides $A_xB_yO_z(s)$
  - Composition and structure available experimentally
  - Energy calculated
- Oxygen is taken from water (and hydrogen molecule)

Formation energy:

$$\Delta E = ABO_3(s) - \min \left( c_1 A(s) + c_2 B(s) + c_3 A_xO_y(s) + c_4 B_xO_y(s) + c_5 O \right)$$

$$c_1 + c_3 = 1 \ , \quad c_2 + c_4 = 1 \ , \quad c_3 + c_4 + c_5 = 3$$

→ Solved by linear programming.
Double perovskite example: $\text{BaHfO}_3$, $\text{SrSnO}_3 \rightarrow \text{BaHfSrSnO}_6$

Bandgaps: 6.6 eV, 6.3 eV, 3.0 eV

Conduction band edge:

- $\text{BaHfO}_3$
- $\text{BaHfSrSnO}_6$
- $\text{SrSnO}_3$

Hf(d)-Sn(p) hybridization $\rightarrow$ increased bandgap
Fitting to individual data sets

Training conflict between molecules and solids remain.
Transparent protecting shield – photoanode

$E_{\text{form}} < 0.2$

$E_{\text{gap}} > 3 \text{ eV}$
Difference between the double perovskite bandgap and the average gap coming from the two constituent cubic perovskites

New “design rules”:

- Double perovskite has average of perovskite gaps
- But, for B1 p-metal and B2 d-metal gap is significantly increased (but typically > 4 eV)

- B1-ion(d) - B2-ion(p) hybridization -> increased bandgap
OER overpotentials

\[ G^{OER} = \text{Max}[\Delta G_1^0, \Delta G_2^0, \Delta G_3^0, \Delta G_4^0] \]

Joel Varley, Monica Garcia-Mota, Jens K. Nørskov
Towards bandgap engineering: double perovskites

- 152 new materials for one-photon water splitting
- 100 new materials for two-photon water splitting

Combinations of metals with large-gap semiconductors?
Regularization

Smoothness – Tikhonov reg.
Improves transferability

Bootstrapping

Bootstrapping is used to estimate the dependency of the fit on the database bootstrapping can be used. (Alternative to cross validation).

Consider a database with N points. A bootstrap sample is obtained by selecting N points from the database (with possible repetition)

Averaging over bootstrap samples gives information about both bias and variance

Estimated Prediction Error – 0.632-rule

\[ EPE = 0.368 \text{err} + 0.632 \text{Err}_1 \]

\[ \text{err} = \frac{1}{N} \sum_i (y_i^s - y_i)^2 \]

\[ \text{Err}_1 = \frac{1}{N} \sum_i \frac{1}{\text{Nos}:i \notin s} \sum_{s:i \notin s} (y_i^s - y_i)^2 \]
Cluster analysis: dendrograms

Dendrogram: cluster analysis based on distance

Linkage criterion: cut-off distance $d$; two data points belong to the same cluster if there is a chain of data points with distances less than $d$ which connect them.

Distance between two A-ions:

$$d(A_1, A_2) = \frac{1}{N_B} \sum_B \left( E_{A_1 BO_3}^{gap} - E_{A_2 BO_3}^{gap} \right)^2$$

Measures similarity between rows
Regularization using bootstrap

Smoothness – Tikhonov reg.
  Improves transferability

Bootstrapping
  Estimates the dependency of the fit on the database: Bias and variance
  Estimated Prediction Error – 0.632-rule

\[
C_{\text{reg}} = \omega^2 \int \left( \frac{d^2 F_x(s(t))}{dt^2} \right)^2 dt
\]

Estimated Prediction Error (EPE)
The data sets

- G3/99 Molecular formation energies (223 molecules)
- RE42 Molecular reaction energies (42 reactions)
- DBH24/08 Molecular reaction barriers (12 forward and backward barriers)
- S22x5 Non-covalent interactions (22 molecules at 5 distances)
- Sol34Ec Solid cohesive energies (34 solids)
- Sol27Ec Solid cohesive energies (27 cubic solids)
- Sol27LC Lattice constants (27 cubic solids)
- CE27 Chemisorption energies (27 systems)
Avoid the tail regions where one data set is done extremely well and another very poorly.

One solution: Minimize product of relative costs.