Consistent Gaussian Basis Sets for Solid State Calculations

Michael F. Peintinger

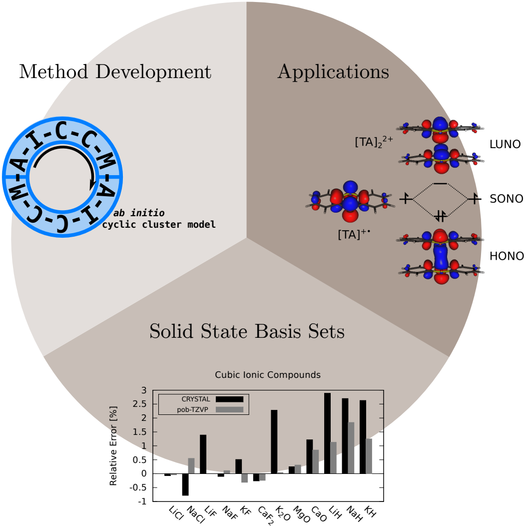
Mulliken Center for Theoretical Chemistry
University of Bonn

2013/02/07

Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
    - Molecular versus solid state basis sets
    - Optimizing basis sets for solid state calculations
    - Implementation of the optimization procedure
    - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
    - pob-TZVP and pob-TZVPP
    - pob-DZVP, pob-DZVPP (under development)
    - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

Research interest

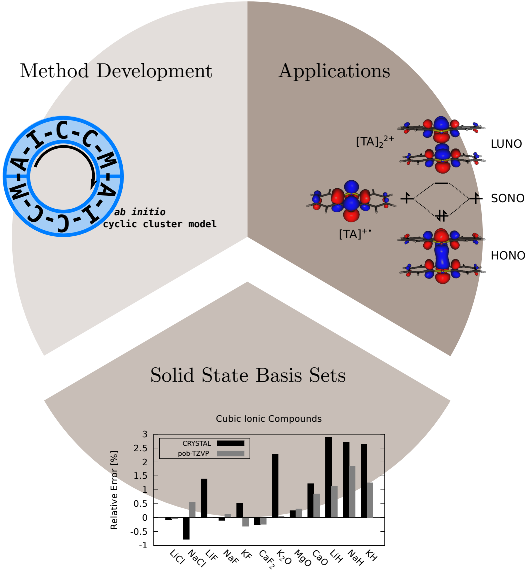


Method Development

Applications

Solid State Basis Sets

Research interest



Method Development

Applications

Solid State Basis Sets

Motivation

- In quantum chemical calculations the choice of basis set has a huge impact on the quality of the results
- For *molecular calculations*, there is a huge variety of standard basis sets available
- In solid state quantum chemistry, basis sets are often optimized for the system of interest
- Optimized basis not portable to systems that they were not optimized for

Motivation

- In quantum chemical calculations the choice of basis set has a huge impact on the quality of the results
- For *molecular calculations*, there is a huge variety of standard basis sets available

- In solid state quantum chemistry, basis sets are often optimized for the system of interest
- Optimized basis not portable to systems that they were not optimized for

Plane waves versus Gaussian basis sets

Plane waves

- ☹ Large number of plane waves needed
- ☹ Must use pseudo-potentials
- ☹ Loss of chemical insight
- ☺ Numerically efficient (fast Fourier transform)
- ☺ One parameter controls basis set size (kinetic energy cutoff)
- ☺ No basis set superposition error (BSSE)
- ☹ Pulay stress

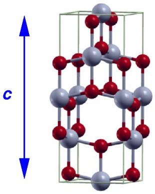
Gaussian basis sets

- ☺ Good results for small basis sets
- ☺ All-electrons calculations and PP
- ☺ Correspond to chemical insight
- ☺ Computationally efficient (multi-centre integrals)
- ☹ Systematic improvement not straightforward
- ☹ Linear dependencies, over-completeness problem
- ☹ Basis set superposition error (BSSE)

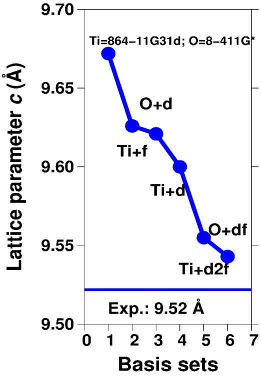
LCAO Basis Set Dependence

- Geometries are expected to be basis set independent
- Various exceptions, e.g. lattice parameter **c** of anatase (large effect of diffuse and polarization functions)

Gadaczek, Esch, Bredow, in preparation



anatase unit cell



Other work:

Method	*c*
HF	9.688
PBE	9.867
B3LYP	9.805
PBE0	9.704
LDA	9.495
Expt.	9.522

Labat et al., JPC 126 (2007) 154703

What are we aiming for on the long run?

- Reliable and consistent basis sets for solid state calculations
- Various quality levels to choose from
- Estimate the error due to incompleteness of the basis

Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
    - Molecular versus solid state basis sets
    - Optimizing basis sets for solid state calculations
    - Implementation of the optimization procedure
    - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
    - pob-TZVP and pob-TZVPP
    - pob-DZVP, pob-DZVPP (under development)
    - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

Methods for the generation of basis sets

Efficiency versus accuracy

The two conflicting goals for basis sets, **fast** and **accurate** calculations, have led to the generation of many different basis sets.

Different approaches to the design of basis sets

- Direct fitting to Slater-Type Orbitals (STOs), e.g. STO-3G
- Variationally minimizing HF or DFT energies, e.g. TZVP
- Minimizing post-HF energies, e.g. cc-pVDZ
- Fitting to properties, e.g. pc-3

Methods for the generation of basis sets

Efficiency versus accuracy

The two conflicting goals for basis sets, **fast** and **accurate** calculations, have led to the generation of many different basis sets.

Different approaches to the design of basis sets

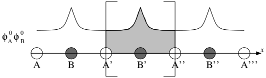
- Direct fitting to Slater-Type Orbitals (STOs), e.g. STO-3G
- Variationally minimizing HF or DFT energies, e.g. TZVP
- Minimizing post-HF energies, e.g. cc-pVDZ
- Fitting to properties, e.g. pc-3

universitätbonn

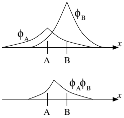
Linear Dependence

- Condition for linear dependence: $\sum_{b \neq a} S_{ab} \geqslant 1$
- Overlap between two Bloch functions (BF)

$$S^0_{ab} = \int_0^T \phi^0_a(x)\phi^0_b(x)\,dx$$



- Overlap between two AOs



- In general: $S^0_{ab}(BF) > S_{ab}(AO)$

Using molecular basis sets for solid state calculations

Although molecular basis sets can in principle be used in solid state calculations if exponents below a certain threshold are cut off, the following problems arise:

- Basis sets are usually optimized in atomic calculations.
- Cutting off below a certain threshold leads to basis sets which are inconsistent regarding size
- By switching to another problem (e.g. adsorption), the basis sets might become unusable again

Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
   - Molecular versus solid state basis sets
   - Optimizing basis sets for solid state calculations
   - Implementation of the optimization procedure
   - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
   - pob-TZVP and pob-TZVPP
   - pob-DZVP, pob-DZVPP (under development)
   - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Strategy

1. Select established molecular basis set (e.g. def2-TZVP)
2. Cut off valence functions with exponents below a certain threshold (e.g. > 0.2)
3. Add basis functions until desired quality level is reached, exponents are set to $\frac{1}{3}$ to $\frac{1}{2}$ of the preceeding function
4. Reoptimize exponents and contraction coefficients of the valence shells in various binary compounds representing different bonding situations by variationally minimizing the total energy with a hybrid method (e.g. PW1PW)
5. Cross-check the resulting basis sets in all test systems
6. Select the overall best performing (and stable) basis set and reoptimize exponents and contraction coefficients one at a time until good results in all test systems are achieved

Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
    - Molecular versus solid state basis sets
    - Optimizing basis sets for solid state calculations
    - Implementation of the optimization procedure
    - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
    - pob-TZVP and pob-TZVPP
    - pob-DZVP, pob-DZVPP (under development)
    - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

ROOT, Minuit and PyMinuit

What is ROOT?

The ROOT system provides a set of OO frameworks with all the functionality needed to handle and analyze large amounts of data in a very efficient way and is widely used in particle physics

What is Minuit?

Minuit has been the standard package for minimizing general N-dimensional functions and it features a robust set of algorithms for optimizing the search, correcting mistakes, and measuring non-linear error bounds.

What is PyMinuit?

PyMinuit is an extension module for Python that passes low-level Minuit functionality to Python functions.

In practice...

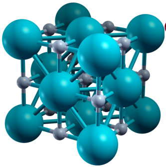
```
K2S F m3m ICSD 641321 pob
CRYSTAL
…
END
19 11
…
0 0 1 1 1
  ${a1=0.90384027}  1.000000000
0 0 1 0 1
  ${a2=0.28693087}  1.000000000
0 0 1 0 1
  ${a3=0.11669884}  1.000000000
0 2 6 6 1
  728.18449873      0.0026150689792
  172.13265061      0.020673630835
…
```

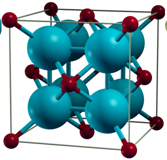
- User writes CRYSTAL input with exponents and coefficients marked as variables,
- runs BSOpt script (written in Python),
- which calls CRYSTAL and parses output for total energy
- and uses PyMinuit to minimize N-dimensional problem.

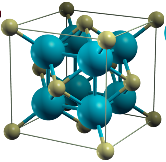
Outline

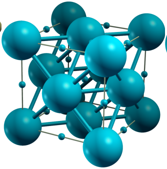
1. Introduction
2. Consistent optimization of Gaussian basis sets
    - Molecular versus solid state basis sets
    - Optimizing basis sets for solid state calculations
    - Implementation of the optimization procedure
    - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
    - pob-TZVP and pob-TZVPP
    - pob-DZVP, pob-DZVPP (under development)
    - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

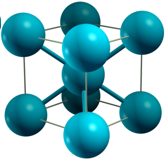
Optimization of Potassium basis set


(a) KF


(b) K_2O

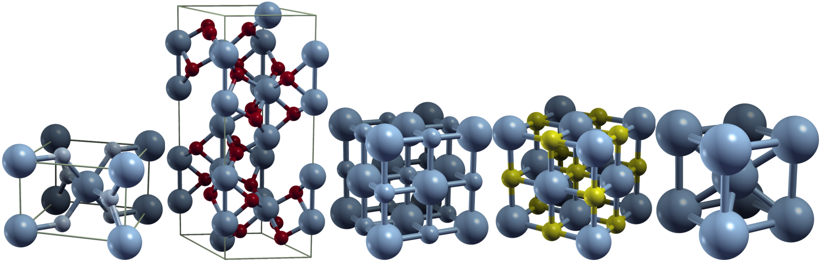

(c) K_2S


(d) KH


(e) K

Compound	Experiment	pob-TZVP	
KF	5.347	5.364	
K_2O	6.436	6.434	
KH	5.704	5.633	
K_2S	7.407	7.350	
K	5.247	4.570	5.141

Optimization of Vanadium basis set



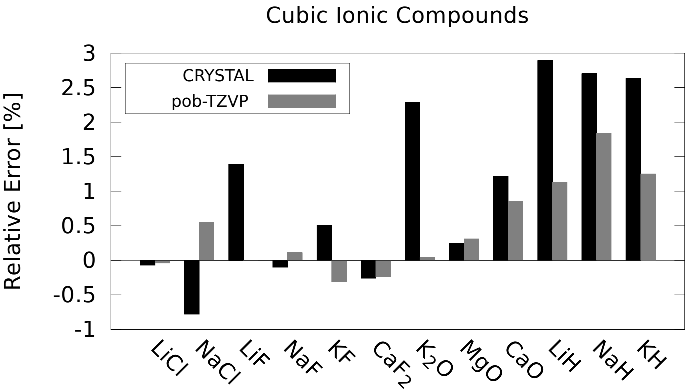
(a) VF_2 (b) V_2O_3 (c) VN (d) VC (e) V

Compound	Experiment	pob-TZVP
VF_2	4.806, 3.237	4.805, 3.225
V_2O_3	4.952, 14.002	5.038, 13.835
VN	4.137	4.078
VC	4.163	4.152
V	3.024	2.946

Outline

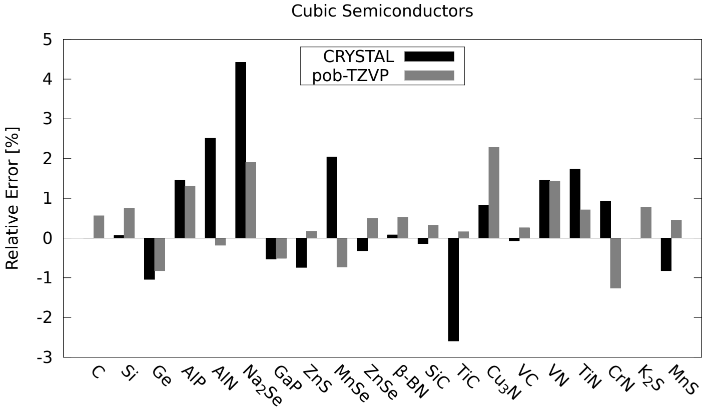
1. Introduction
2. Consistent optimization of Gaussian basis sets
   - Molecular versus solid state basis sets
   - Optimizing basis sets for solid state calculations
   - Implementation of the optimization procedure
   - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
   - pob-TZVP and pob-TZVPP
   - pob-DZVP, pob-DZVPP (under development)
   - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

Cubic ionic compounds



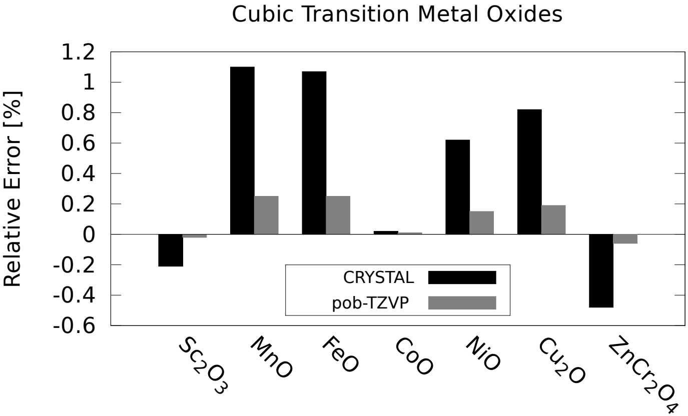
Relative error in the lattice constants of cubic compounds with respect to experimental values. Maximum error : CRYSTAL 2.9%, pob-TZVP 1.8%; Mean error: CRYSTAL 1.3%, pob-TZVP 0.7%

Cubic semiconducting compounds



Relative error in the lattice constants of cubic semiconductors with respect to experimental values. Maximum error : CRYSTAL 4.4%, pob-TZVP 3.1%; Mean error: CRYSTAL 3.1%, pob-TZVP 1.0%

Cubic transition metal compounds



Relative error in the lattice constants of cubic transition metal oxides with respect to experimental values. Maximum error : CRYSTAL 2.9%, pob-TZVP 1.8%; Mean error: CRYSTAL 1.3%, pob-TZVP 0.7%

Additional polarization functions

Experimental and calculated lattice parameters of selected ionic solid cubic compounds with CRYSTAL standard, pob-TZVP and pob-TZVPP basis sets. Lengths are given in Å. Maximum error: 2.7% (CRYSTAL), 1.4% (pob-TZVP), 0.6% (pob-TZVPP); Mean error: 1.6% (CRYSTAL), 1.1% (pob-TZVP), 0.2% (pob-TZVPP)

Compound	Experiment	CRYSTAL	pob-TZVP	pob-TZVPP
	a	**a**	**a**	**a**
NaCl	5.640	5.684	5.609	5.640
CaO	4.811	4.752	4.770	4.773
NaH	4.890	4.758	4.800	4.812
KH	5.704	5.554	5.633	5.642
K_2S	7.407	7.407	7.350	7.370

Thermochemistry

Experimental and calculated atomization enthalpies of selected ionic solid cubic compounds with CRYSTAL standard and pob-TZVP basis sets. Energies are given in $\frac{kJ}{mol}$.

Compound	Experiment	CRYSTAL	pob-TZVP
LiCl	685	663	654
NaCl	640	614	614
LiF	849	834	823
NaF	762	733	738
KF	737	705	692
CaF_2	1562	1531	1540
K_2O	790	701	683
MgO	994	971	970
CaO	1062	1030	1034
LiH	518	471	457
NaH	382	374	379
KH	365	349	333

Metals

Experimental and calculated lattice constants of selected cubic metals. A hyphen indicates that the optimization of the lattice constant failed. The value in brackets is the orbital exponent of the added function. Lengths are given in Å.

Metal	Experiment	CRYSTAL	pob-TZVP	pob-TZVP+s	shell (Exponent)
Li	4.404	3.871	4.390	4.444	s (0.050)
Na	4.291	--	4.269		
K	5.247	--	4.570	5.141	s (0.033)
Ca	4.486	--	4.035	4.324	s (0.067)
Sc	4.541	--	4.428		
V	3.024	2.637	2.946		
Cr	3.680	3.233	3.473		
Fe	2.867	--	2.746		
Ni	3.524	3.171	3.423		
Cu	3.615	3.363	3.506		
Al	4.050	3.494	3.731	4.092	p (0.05)
Ni_3Al	3.550	3.105	3.434		

Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
   - Molecular versus solid state basis sets
   - Optimizing basis sets for solid state calculations
   - Implementation of the optimization procedure
   - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
   - pob-TZVP and pob-TZVPP
   - pob-DZVP, pob-DZVPP (under development)
   - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

A glimpse on the pob-DZVP/PP basis sets

Experimental and calculated lattice parameters of selected ionic solid cubic compounds with CRYSTAL standard and pob basis sets. Lengths are given in Å.

Comp.	Exp.	CRYSTAL	pob-TZVP	**pob-DZVP**	**pob-DZVPP**
	a	**a**	**a**	**a**	**a**
NaCl	5.640	5.609	5.609	5.640	5.640
K_2O	6.436	6.289	6.434	6.434	6.433
MgO	4.217	4.207	4.204	4.200	4.197
TiC	4.328	4.443	4.321	4.294	--

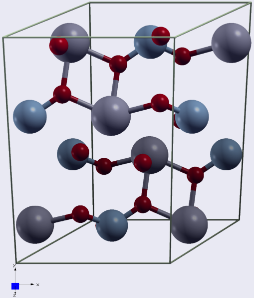
Outline

1. Introduction
2. Consistent optimization of Gaussian basis sets
   - Molecular versus solid state basis sets
   - Optimizing basis sets for solid state calculations
   - Implementation of the optimization procedure
   - K and V as examples for basis optimizations
3. Results obtained with pob basis sets
   - pob-TZVP and pob-TZVPP
   - pob-DZVP, pob-DZVPP (under development)
   - pob-ECP-DZVP,pob-ECP-DZVPP and pob-ECP-TZVP (under development)
4. Thoughts on the complete basis set limit for solid state calculations

$BiVO_4$ and $InVO_4$

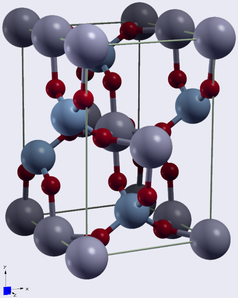
- Promising photocatalysts for hydrogen generation
- Large supercells, defects and surfaces

$BiVO_4$



$BiVO_4$

$InVO_4$



$InVO_4$

Structure and band gap

Preliminary calculations

- Structure is ok (CRYSTAL09, PW1PW, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP)
- Band gap is bad: Δ+1.3eV resp. Δ+1.9eV (Exp. 2.4 eV resp. 3.2 eV)

This is not a basis set problem!

- Literature (VASP, PBE) : 2.2 eV

  Walsh, A. et. al., Chem. Mat. 2009 21 (3), 547-551.
- CRYSTAL09 (PBE, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP) : 2.2 eV

$\Rightarrow$ Reduce exchange empirically

Structure and band gap

Preliminary calculations

- Structure is ok (CRYSTAL09, PW1PW, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP)
- Band gap is bad: Δ+1.3eV resp. Δ+1.9eV (Exp. 2.4 eV resp. 3.2 eV)

This is not a basis set problem!

- Literature (VASP, PBE) : 2.2 eV

  Walsh, A. et. al., Chem. Mat. 2009 21 (3), 547-551.

- CRYSTAL09 (PBE, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP) : 2.2 eV

⇒ Reduce exchange empirically

Structure and band gap

Preliminary calculations

- Structure is ok (CRYSTAL09, PW1PW, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP)
- Band gap is bad: Δ+1.3eV resp. Δ+1.9eV (Exp. 2.4 eV resp. 3.2 eV)

This is not a basis set problem!

- Literature (VASP, PBE) : 2.2 eV

  Walsh, A. et. al., Chem. Mat. 2009 21 (3), 547-551.

- CRYSTAL09 (PBE, Bi/In ECP-DZVPP, O-DZVPP, V-DZVP) : 2.2 eV

⇒ Reduce exchange empirically

Complete basis set limit

Complete basis set limit (CBS)

The complete basis set (CBS) limit is an extrapolated estimate of a result obtained using an infinitely large (complete) basis set.

- In principle removes any error due to the linear combination of atomic orbitals (LCAO) approximation
- Remaining disagreement with experiment is due to other approximations (treatment of electron correlation)
- For most properties the CCSD(T) in CBS limit results can be regarded as numerically exact for all practical purposes
- Potential to end the almost endless discussion about Gaussian basis sets in solid state applications

Complete basis set limit

Complete basis set limit (CBS)

The complete basis set (CBS) limit is an extrapolated estimate of a result obtained using an infinitely large (complete) basis set.

- In principle removes any error due to the linear combination of atomic orbitals (LCAO) approximation
- Remaining disagreement with experiment is due to other approximations (treatment of electron correlation)
- For most properties the CCSD(T) in CBS limit results can be regarded as numerically exact for all practical purposes
- Potential to end the almost endless discussion about Gaussian basis sets in solid state applications

Requirements

- Minimum of three separate calculations with increasingly larger basis sets
- Extrapolation works only with basis sets designed specifically for the task
  $\Rightarrow$ Correlation- or polarization-consistent basis sets, e.g. cc-pVxZ or pc-n

- Adjustment of cc-pVxZ basis sets for periodic case?
- Can our consistent solid state basis sets be used?

Requirements

- Minimum of three separate calculations with increasingly larger basis sets
- Extrapolation works only with basis sets designed specifically for the task
  ⇒ Correlation- or polarization-consistent basis sets, e.g. cc-pVxZ or pc-n

- Adjustment of cc-pVxZ basis sets for periodic case?
- Can our consistent solid state basis sets be used?

Correlation- or polarization-consistent basis sets

- Most popular are those developed by Dunning and coworkers
    - Double-zeta: cc-pVDZ
    - Triple-zeta: cc-pVTZ
    - Quadruple-zeta: cc-pVQZ
    - Quintuple-zeta: cc-pV5Z
    - Augmented with diffuse functions: aug-cc-pVDZ, ...
- Most popular polarization consistent basis sets are those developed by Jensen and coworkers
    - pc-n (n=1, 2, ...)
    - n indicates the largest increment in angular momentum above the valence maximum used in the basis set

Extrapolation technique

1. Property Y of interest (e.g. a relative energy, a frequency, or a bond length) is computed at a given level of theory (e.g. HF, DFT or Hybrid) using at least three basis sets (e.g. cc-VDZ, cc-VTZ, and cc-VQZ)
2. These data points are then fitted to an equation

Extrapolation technique

Popular fitting equations

- Exponential form:

$$E(L_{\text{max}}) = E_{\text{CBS}} + \alpha e^{-\beta L_{\text{max}}} \tag{1}$$

- Power form:

$$E(L_{\text{max}}) = E_{\text{CBS}} + \alpha L_{\text{max}}^{-3} \tag{2}$$

E_{CBS} : estimated complete basis set limit
DZ: $L_{\text{max}} = 2$, TZ: $L_{\text{max}} = 3$, ...

Summary

- There is a need for consistent standard Gaussian basis sets for solid state calculations.

- Outlook
    - Continue development of basis sets at various quality levels
    - Enlarge the test set and extend the tested properties
    - Complete basis set limit for solid state calculations?

Thanks to...

Prof. Dr. Thomas Bredow

Daniel Vilela Oliveira
Christoph Reimann

Dr. Jonas Baltrusaitis (University of Twente)

dfg·sfb 813
Chemie an Spinzentren
SFB 813 der Deutschen Forschungsgemeinschaft

SFB813 - Chemistry at Spin Centers

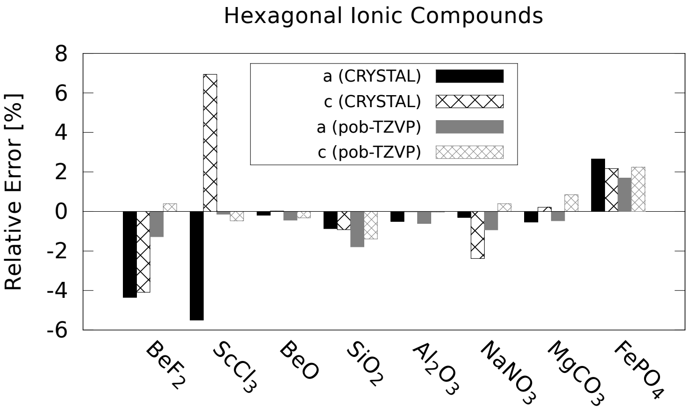
universitätbonn


POSTAMT
HARIBO
D-53129 Bonn

FIN

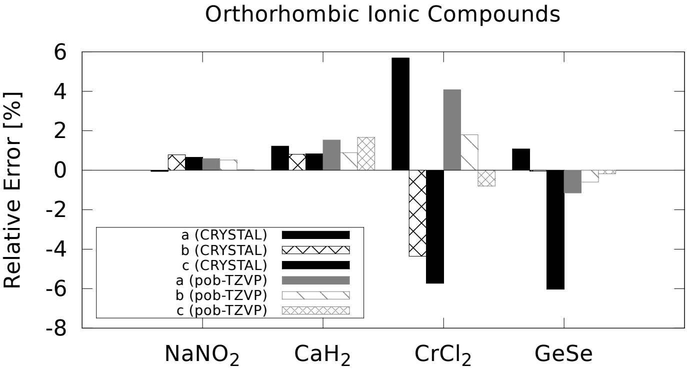
Thank you!

Hexagonal ionic compounds



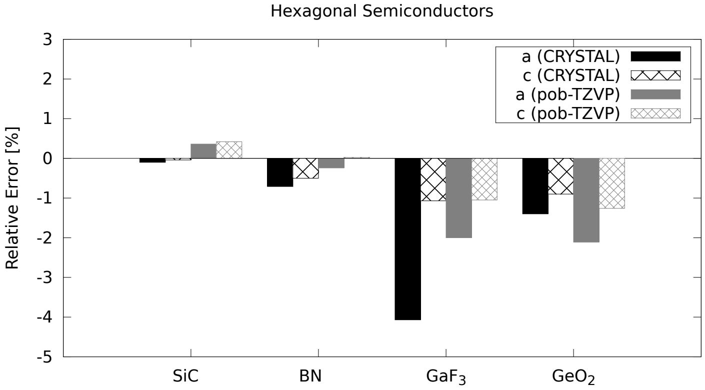
Relative error in the lattice constants of hexagonal compounds with respect to experimental values. Maximum error: CRYSTAL 12.5%, pob-TZVP 3.9%; Mean error: CRYSTAL 2.0%, pob-TZVP 0.8%

Orthorhombic ionic compounds



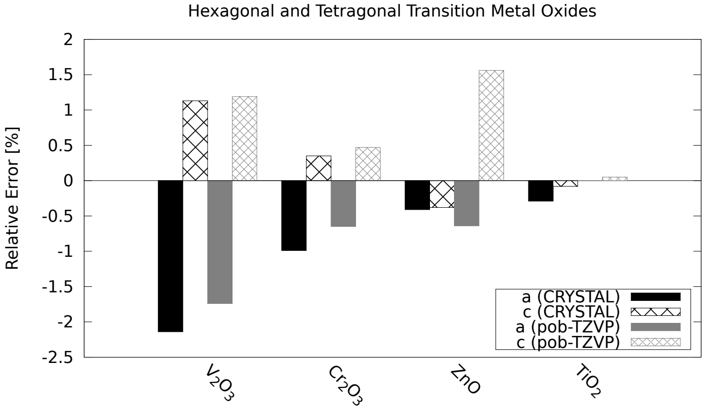
Relative error in the lattice constants of orthorhombic compounds with respect to experimental values. Maximum error: CRYSTAL 15.8%, pob-TZVP 6.7%; Mean error: CRYSTAL 2.3%, pob-TZVP 1.2%.

Hexagonal semiconducting compounds



Relative error in the lattice constants of hexagonal compounds with respect to experiment Maximum error: CRYSTAL 30.6%, pob-TZVP 9.8%; Mean error: CRYSTAL 3.1%, pob-TZVP 1.0%.

Tetragonal transition metal compounds



Relative error in the lattice constants of hexagonal and tetragonal transition metal oxides with respect to experiment Maximum error: CRYSTAL 4.6%, pob-TZVP 4.0%; Mean error (sum): CRYSTAL 1.3%, pob-TZVP 1.1%.