TaskBlaster:

An open Python framework for FAIR and scalable workflows

Kristian Thygesen & Ask Larsen

Technical University of Denmark

thygesen@fysik.dtu.dk

Outline

- Automated workflows for computational materials science
- Constructing a computational database of 2D materials
- Can AI help to generate new stable crystals?
- Conclusions

Automated workflows (or Big Data drivers)

- Computational materials design requires **massive amounts of calculations**
- This calls for **automated workflow solutions** that
  - Submit calculations
  - Receive and store results
  - Keep track of the data provenance

[PhD student performing high-throughput calculations manually]

Courtesy of G. Pizzi

The basic philosophy behind the ASR

The basic philosophy behind the ASR

High-throughput projects are iterative and data must be maintained over time.
Workflows should be easy to adjust and rerun without compromising data provenance.

Components of the ASR workflow framework

Components of the ASR workflow framework

Atomic Simulation Environment (ASE)

- Python scripting environment for working with atoms and materials simulations
- Historically the first of its kind; first paper by Bahn & Jacobsen in 2002.
- Community driven
- **Interface to 30+ atomistic simulation codes**

Highly active user/developer community

>324 contributers

A. H. Larsen *et al., J. Phys.: Condens. Mat.* **29** 273002 (2017)
S. R. Bahn and K. W. Jacobsen, *Comput. Sci. Eng.* **4**, 56 (2002)

Components of the ASR workflow framework

The GPAW code

Mortensen and Jacobsen, PRB B **71**, 035109 (2005) ← **First paper**

Enkovaara *et al.* J. Phys.:Cond. Mat. **22** (2010) ← **Review article**

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

- ❑ Projector augmented wave (PAW)
- ❑ Three different basis sets (real space, plane waves, LCAO)
- ❑ Efficient parallelization (good scalability up to > 100.000 CPU cores)
- ❑ Time-dependent DFT (linear response + time propagation)
- ❑ Many-body perturbation theory (RPA, GW(Γ), BSE)
- ❑ Phonons and electron-phonon coupling
- ❑ Quantum electron transport
- ❑ QM/MM
- ❑ Well integrated with the Atomic Simulation Environment (ASE) and Recipes (ASR)

Components of the ASR workflow framework

TaskBlaster: The ASR backend

The possible states of a Task

Static and dynamic tasks

- ○ Static task
- ⊕ Dynamic task: Can create new tasks depending on the result of other tasks

Components of the ASR workflow framework

Task execution using “workers”

- Workers are computational processes to run tasks
- Each worker picks up available tasks and runs them in sequence
- Tasks can specify which kind of worker may pick them up; different kinds of worker can have different parallelizations
- Workers can be submitted using MyQueue
- Worker configuration can be specified at project level

Scaling of TaskBlaster

Ask Larsen

Jens Mortensen

Mikael Kuisma

Tara Boland

Julian Heske

+ assistance from CSC

Demonstrated scaling of TaskBlaster to the entire LUMI supercomputer (g-partition).

Running ca. 14000 GPAW-GPU calculations concurrently.

Not trivial! Experience used to develop a LUMI-compatible version of ASR

- Approx. 10k GPUs
- 375 petaflops (or 1.5 mio. laptops)
- #3 in the world
- Operated by CSC, Finland

C2DB

The open Computational 2D Materials Database

Known layered bulk compounds (~600)

M. Ashton *et al.* PRL **118**, 106101 (2017)
Mounet *et al*. Nature Nano. **13**, 246 (2018)

Top-down

Bottom-up

2D crystal structure prototypes

Combinatorial lattice decoration

Hypothetical 2D materials

C2DB The open Computational 2D Materials Database

Known layered bulk compounds (~600)

M. Ashton *et al.* PRL **118**, 106101 (2017)
Mounet *et al*. Nature Nano. **13**, 246 (2018)

Top-down

Bottom-up

2D crystal structure prototypes

Combinatorial lattice decoration

Hypothetical 2D materials

C2DB workflow

Structure relaxation (symmetry constrained)

Dimensionality analysis: Exactly one 2D cluster ?

no → not new / 2D

yes

Duplicate check: Unique material ∉ C2DB ?

no → not new / 2D

yes

Prototype classification: "Stoichiometry – SG – Wykoff"

Magnetic classification: $|\mu_{atom}| < 0.1$? , $|\mu_{tot}| < 0.1$?

Formation and hull energy (ΔH_{form} and ΔH_{hull})

$\Delta H_{form} < 0$?

no → thermodynamically unstable

yes

Stiffness tensor (**C**)

min{eig(**C**)} > 0 ?

no → dynamically unstable

yes

Dynamical matrix (**D**)

min{eig(**D**)} > 0 ?

no → dynamically unstable

yes

Property workflow (see Table 1)

- Thermodynamic
- Elastic
- Electronic
- Magnetic
- Topological
- …

monolayers: C2DB

multilayers: BiDB

point defects: QPOD

C2DB

The Computational 2D Materials Database

Components of the ASR workflow framework

Computational 2D materials database

C2DB
The Computational 2D Materials Database

http://c2db.fysik.dtu.dk

MoS2

Stoichiometry (A, AB2, A2B3, ...) :

Material class : All

Dynamically stable (phonons) : Yes

Dynamically stable (stiffness) : Yes

Thermodynamic stability : Medium - High

Is magnetic : All

Band gap range [eV] : - PBE

Help with constructing advanced search queries ...
Toggle list of keys ...

Displaying rows 1-10 out of 10 (direct link)

Rows: 25 Add Column

Formula	Material class	Space group	Magnetic	Heat of formation	Band gap	Crystal type
Mo2I2S2	-	Pmmn	True	-0.256	0.000	ABC-59-ab
Mo2S4	TMDC-T'	P2_1/m	False	-0.711	0.051	AB2-11-e
Mo2W2S8	TMDC-alloy	Pc	False	-0.866	1.532	ABC4-7-a
MoS2	-	P-4m2	False	-0.306	0.000	AB2-115-dg
Cr3MoS8	TMDC-alloy	P1	False	-0.710	0.899	AB3C8-1-a
Mo3WS8	TMDC-alloy	P1	False	-0.880	1.559	AB3C8-1-a
MoS2	TMDC-H	P-6m2	False	-0.921	1.580	AB2-187-bi
Cr2Mo2S8	TMDC-alloy	Pma2	False	-0.761	1.014	ABC4-28-acd
CrMo3S8	TMDC-alloy	Pm	False	-0.824	1.184	AB3C8-6-abc
MoW3S8	TMDC-alloy	P1	False	-0.851	1.532	AB3C8-1-a

Computational 2D materials database

C2DB
The Computational 2D Materials Database

http://c2db.fysik.dtu.dk

MoS2

Stoichiometry (A, AB2, A2B3, ...) :

Material class : All

Dynamically stable (phonons) : Yes

Dynamically stable (stiffness) : Yes

Thermodynamic stability : Medium - High

Is magnetic : All

Band gap range [eV] : - PBE

Help with constructing advanced search queries ...
Toggle list of keys ...

Displaying rows 1-10 out of 10 (direct link)

Rows: 25 Add Column

Formula	Material class	Space group	Magnetic	Heat of formation	Band gap	Crystal type
Mo2I2S2	-	Pmmn	True	-0.256	0.000	ABC-59-ab
Mo2S4	TMDC-T'	P2_1/m	False	-0.711	0.051	AB2-11-e
Mo2W2S8	TMDC-alloy	Pc	False	-0.866	1.532	ABC4-7-a
MoS2	-	P-4m2	False	-0.306	0.000	AB2-115-dg
Cr3MoS8	TMDC-alloy	P1	False	-0.710	0.899	AB3C8-1-a
Mo3WS8	TMDC-alloy	P1	False	-0.880	1.559	AB3C8-1-a
MoS2	TMDC-H	P-6m2	False	-0.921	1.580	AB2-187-bi
Cr2Mo2S8	TMDC-alloy	Pma2	False	-0.761	1.014	ABC4-28-acd
CrMo3S8	TMDC-alloy	Pm	False	-0.824	1.184	AB3C8-6-abc
MoW3S8	TMDC-alloy	P1	False	-0.851	1.532	AB3C8-1-a

MoS_2

Download raw data | Browse raw data

C2DB
The Computational 2D Materials Database

http://c2db.fysik.dtu.dk

Summary

Structure info	Value
Crystal type	AB2-187-bi
Material class	TMDC-H
Space group	P-6m2
Space group number	187
Point group	-6m2
Related ICSD id	38401
Related COD id	9007661
Reported DOI	10.1103/PhysRevLett.105.136805

Unit cell

Download

Axis	x (Å)	y (Å)	z (Å)	Periodic
1	3.184	0.000	0.000	True
2	-1.592	2.757	0.000	True
3	0.000	0.000	18.127	False

Lengths (Å):	3.184	3.184	18.127
Angles (°):	90.000	90.000	120.000

Stability

Thermodynamic	HIGH
Dynamical (phonons)	HIGH
Dynamical (stiffness)	HIGH

Electronic properties

Magnetic	False
Band gap (PBE)	1.58 eV
Band gap (HSE)	2.09 eV
Band gap (G0W0)	2.53 eV

- Thermodynamic stability
- Stiffness tensor
- Phonons
- Basic electronic properties (PBE)
- Electronic band structure (PBE)
- Projected band structure and DOS (PBE)
- Effective masses (PBE)
- Electronic band structure (HSE)
- Electronic band structure (G0W0)
- Born charges
- Optical polarizability (RPA)
- Infrared polarizability (RPA)
- Raman spectrum
- Optical absorption (BSE and RPA)
- Bader charges
- Piezoelectric tensor

MoS_2

Download raw data | Browse raw data

C2DB
The Computational 2D Materials Database

http://c2db.fysik.dtu.dk

Summary

Structure info	Value
Crystal type	AB2-187-bi
Material class	TMDC-H
Space group	P-6m2
Space group number	187
Point group	-6m2
Related ICSD id	38401
Related COD id	9007661
Reported DOI	10.1103/PhysRevLett.105.136805

Unit cell

Download

Axis	x (Å)	y (Å)	z (Å)	Periodic
1	3.184	0.000	0.000	True
2	-1.592	2.757	0.000	True
3	0.000	0.000	18.127	False

Lengths (Å):	3.184	3.184	18.127
Angles (°):	90.000	90.000	120.000

Stability

Thermodynamic	HIGH
Dynamical (phonons)	HIGH
Dynamical (stiffness)	HIGH

Electronic properties

Magnetic	False
Band gap (PBE)	1.58 eV
Band gap (HSE)	2.09 eV
Band gap (G0W0)	2.53 eV

Thermodynamic stability

Stiffness tensor

Phonons

General panel information

The effective mass tensor represents the second derivative of the band energy w.r.t. wave vector at a band extremum. The effective masses of the valence bands (VB) and conduction bands (CB) are obtained as the eigenvalues of the mass tensor. The latter is determined by fitting a 2nd order polynomium to the band energies on a fine k-point mesh around the band extrema. Spin-orbit interactions are included. The "parabolicity" of the band is quantified by the mean absolute relative error (MARE) of the fit to the band energy in an energy range of 25 meV.

Relevant article(s):

- S. Haastrup et al. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals, 2D Mater. 5 042002 (2018).

Relevant recipes

This panel contains information calculated with the following ASR Recipes:

- asr.emasses@validate

Optical polarizability (RPA)

Infrared polarizability (RPA)

Raman spectrum

Optical absorption (BSE and RPA)

Bader charges

Piezoelectric tensor

http://c2db.fysik.dtu.dk

Electronic band structure (PBE)

Interactive mode

- Thermodynamic stability
- Stiffness tensor
- Phonons
- Basic electronic properties (PBE)
- Electronic band structure (PBE)
- Projected band structure and DOS (PBE)
- Effective masses (PBE)
- Electronic band structure (HSE)
- Electronic band structure (G0W0)
- Born charges
- Optical polarizability (RPA)
- Infrared polarizability (RPA)
- Raman spectrum
- Optical absorption (BSE and RPA)
- Bader charges
- Piezoelectric tensor

http://c2db.fysik.dtu.dk

Electronic band structure (PBE)

Interactive mode

Raman spectrum

Mode	Frequency (1/cm)	Degeneracy
Mode 1	0.	3
Mode 2	283.7	2
Mode 3	383.8	2
Mode 4	407.9	1
Mode 5	473.2	1

- Thermodynamic stability
- Stiffness tensor
- Phonons
- Basic electronic properties (PBE)
- Electronic band structure (PBE)
- Projected band structure and DOS (PBE)
- Effective masses (PBE)
- Electronic band structure (HSE)
- Electronic band structure (G0W0)
- Born charges
- Optical polarizability (RPA)
- Infrared polarizability (RPA)
- Raman spectrum
- Optical absorption (BSE and RPA)
- Bader charges
- Piezoelectric tensor

Computational 2D Materials Database (C2DB)

Ca. 4000 monolayers

Properties

- Formation energy
- Energy above convex hull
- Dynamical stability
- Stiffness tensor
- Phonons
- Electronic band structure
- Orbital-projected DOS
- Spin-projected DOS
- Plasma frequency
- Effective masses
- Topological invariants
- Magnetic moment
- Magnetic anisotropy
- Exchange coupling
- Born charges
- Piezoelectric tensor
- Optical conductivity
- Exciton binding energy
- IR conductivity
- Raman spectrum
- Second harmonics generation
- Shift current
- Spontaneous polarisation

Conclusions

- **Workflows** are essential not only for high-throughput but also to ensure reproducibility of results and FAIR data
- **C2DB:** open access to a highly curated structure/property data set for thousands of 2D materials (point defects and bilayers coming up…)
- **ASR/TaskBlaster/MyQueue:** Python workflow framework with focus on simplicity, modularity, flexibility, and transparency.
  Goal is to achieve simulation code agnostic workflows (challenging!)

Thanks!

Ask Hjorth Larsen

Morten Gjerding

Jens Jørgen Mortensen

Tara Boland

Mikael Kuisma

Peder Lyngby

Jiban Kangsabanik

Nikolaj Knoesgaard

… and the rest of CAMD

Center for Nanostructured Graphene

European Research Council

Danmarks Grundforskningsfond
Danish National Research Foundation

VILLUM FONDEN

novo nordisk fonden

NOMAD

Sponsored by the European Union H2020