### **Computational Screening of Soft Materials** Systems with Application to Nano-**Lubrication Systems**

#### **Peter T. Cummings**

Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN Vanderbilt University Multiscale Modeling and Simulation (MuMS) Center, Nashville, TN



# Motivation - Computational Materials Discovery

- □ Computational discovery of new materials is emerging as new subfield of computational materials science developing out of the Materials Genome Initiative (MGI), e.g.:
  - o The Materials Project
    - https://materialsproject.org
    - Properties of over 80,000 materials and screened 25,000 of these for Li-ion batteries.
      - Based on use of low-level electronic structure methods on crystalline (ordered) materials

#### o Screening of metal-organic frameworks (MOFs)<sup>1</sup>

- 137,953 hypothetical MOF structures generated
  - Each MOF evaluated for methane storage at 35 bar and 298 K using grand canonical Monte Carlo
- ~300 MOFs had higher methane-storage capacity at 35 bar than current world record

#### o Crowdsourced Clean Energy Project database at H

- http://www.molecularspace.org
- Data and analyses on 2.3 million candidate compounds for organic photovoltaics
  - Ab initio methods applied to single molecules, linking structural and electronic properties



[1] Wilmer, et al., Large-scale screening of hypothetical metal–organic frameworks. Nature Chemistry, 4 (2011) 83-89. doi:10.1038/nchem.1192



 $https://www.mgi.gov/sites/default/files/documents/materials\_genome\_initiative-final.pdf and the set of the s$ 





## **Motivation - MGI for Soft Matter**

□ In soft matter, predominant physical behaviors occur at energy scales comparable with  $k_BT$ Quantum aspects generally unimportant

- o Weak dispersion interactions frequently dominate
  - Contrast with hard materials
    - Energy scales large by comparison to k<sub>B</sub>T
    - Relatively easy to predict properties since atoms/molecules are typically in crystalline lattice
- o It can be difficult to predict properties directly from atomic or molecular constituents
  - Soft matter often self-organizes into mesoscopic physical structures that are much larger than microscopic scale yet much smaller than the macroscopic scale
    - Macroscopic properties are result of mesoscopic structure
    - Hard materials typically have no mesoscopic structuring
  - Many steps are needed to gather properties based on a given chemistry and condition





#### The Molecular Simulation and Design Framework

□ Motivation - develop a platform for performing MGI-style screening of soft-matter systems

o Goal:



o The procedures grouped into the arrow can be considered a "cottage industry" that is not always easy to reproduce

- Only a few people have the knowledge to reproduce the specific simulation
  - Procedures are not always well documented/described
  - The properties may depend on the exact details of the procedures
- Ad hoc tools/scripts are not always made available
  - Software may be highly specific and not extensible
  - Software may be poorly written and not be well validated



SCHOOL OF ENGINEERING

MUMS

5

### A Molecular Simulation and Design Framework

Motivation - develop a platform for performing Materials Genome Initiative (MGI) style screening of soft-matter systems



#### A Molecular Simulation and Design Framework

# Motivation - develop a platform for performing Materials Genome Initiative (MGI) style screening of soft-matter systems

- o To do so, need to automate each of these steps and be able to combine them together
  - Model-Integrated Computing (MIC)
  - Developed by Vanderbilt researchers in the Institute for Software Integrated Systems (ISIS)
  - Used in wide variety of applications where scientific/engineering workflows require automation and integration to achieve complex goals
    - Car manufacturing assembly line automation and control (GM)
    - Health records management (NIH)
    - Amphibious infantry fighting vehicle design (DARPA)
  - MIC involves meta-language abstraction
- o MoSDeF is outgrowth of three NSF grants joint with ISIS
  - NSF CDI CBET 1028374, NSF SSI ACI 1047828, NSF SSE ACI 1535150
  - MoSDeF also seeks to address simulation reproducibility



SCHOOL OF ENGINEERING Sztipanovits, J. Karsai, G., "Model-integrated computing," Computer, 30, 110-111 (1997)

#### A Molecular Simulation and Design Framework

□ Why does MoSDeF promote simulation reproducibility and why is this important?

- o In a science-skeptic world, the "reproducibility crisis" in scientific research has emerged as critical issue
  - http://www.nature.com/news/reproducibility-1.17552
- o Reproducibility in computational science is becoming a forefront issue in the computational science community
  - Automation is a key device for reproducibility
    - Sandve, G. K. et al. (2013), Ten simple rules for reproducible computational research, PLOS Comp. Bio. (editorial), Vol. 9(10):1–4, doi: 10.1371/ journal.pcbi.1003285
  - "Reproducible computational research, in which all details of computations—code and data—are made conveniently available to others, is a necessary response to [the credibility] crisis."
    - Donoho, D. et al. (2009), Reproducible research in computational harmonic analysis, Comp. Sci. Eng. 11(1):8–18, doi: 10.1109/MCSE.2009.15
  - GUIs are the enemy of reproducibility
    - Lorena A. Barba, George Washington University; see her blog <u>http://lorenabarba.com/category/blog/</u> and "The hard road to reproducibility", Science, 354(6308):142 (October 2016)
- Scriptability, open-source, complete data dissemination needed to reproduce published simulations are goals inherent to MoSDeF



SCHOOL OF ENGINEERING



### A Molecular Simulation and Design Framework

Motivation - develop a platform for performing Materials Genome Initiative (MGI) style screening of soft-matter systems





# **Missing Functionality in Existing Tools**

#### □ Surface functionalization

- o Need ability to express patterns
- o Expose parameters that tune chemistry

#### □ Flexible force field development and dissemination

- o Few force fields exist for all chemistries we care about
- o Not certain how accurate existing FFs will behave under high shear conditions
- o Collaborators and group members developing for specific systems
- Desire to use different simulation engines (meta-level abstraction)
  - o LAMMPS, HOOMD, Gromacs all currently used in group
- MoSDeF is open source, built on open source components







Nuzzo et al., JACS, 1983, 105

## mBuild: a Hierarchical Molecular Builder

□ Sketch or otherwise create simple components (dashed boxes)

□ Combine operations on components into more complicated structures

- o Our applications (nanotribology, supercapacitors, skin lipid self-assembly) typically involve interfacial systems
- □ Everything is scriptable





https://github.com/mosdef-hub/mbuild.git

### **Example: polymerization**

import mbuild as mb
from mbuild.lib.moieties import CH2

polymer = mb.Compound()

last\_monomer = CH2()
polymer.add(last\_monomer)

for \_ in range(10):
 current\_monomer = mb.clone(last\_monomer)

polymer.add(current\_monomer)
last\_monomer = current\_monomer

□ Above code with additional features is provided as mb.Polymer

SCHOOL OF ENGINEERING



# Exposing Tunable Chemistry as Simple Variables



monolayer = AlkaneMonolayer(chain\_length, pattern, n\_tiles)

SCHOOL OF ENGINEERING C. Klein, J. Sallai, T.J. Jones, C.R. Iacovella, C. McCabe, P.T. Cummings, A Hierarchical, Component Based Approach to Screening Properties of Soft Matter, in: Springer, Singapore, 2016: pp. 79–92. doi:10.1007/978-981-10-1128-3\_5.

## Using mBuild: CER Multilayer







# Atom Typing and Applying Force Fields

#### Need general purpose atom typer

- o Arbitrary chemistry means we cannot rely on templates
- o Needs to be easy to develop new atom types and associated logic
- o Don't want to rely on rigid rule hierarchies for atom type definitions
  - Ordered least to most general

#### Removing gap between logic and parameters

- o Parameters exist in various files
- o Logic often encoded as series of if/else statements deep in source code
- o Why not combine the two?
- □ Similar concept being developed by Open Force Field group
  - o https://github.com/open-forcefield-group/



# Foyer: Applying and Disseminating force fields

#### □ End user should only require

- o Force field file
- o Chemical topology

```
from foyer import Forcefield
import parmed as pmd
untyped_ethane = pmd.load_file('ethane.mol2', structure=True)
oplsaa = Forcefield(forcefield_files='oplsaa.xml')
ethane = oplsaa.apply(untyped_ethane)
# Save to any format supported by ParmEd
ethane.save('ethane.top')
ethane.save('ethane.gro')
```





# Adding Atom Type Definitions to OpenMM XML

#### OpenMM XML file convenient to extend

- o Existing infrastructure
- o Great support for custom functional forms

<ForceField> <AtomTvpes> <Type name="opls\_135" class="CT" element="C" mass="12.01100" def="[C;X4](C)(H)(H)H" desc="alkane CH3"/> <Type name="opls\_140" class="HC" element="H" mass="1.00800" def="H[C;X4]" desc="alkane H"/> </AtomTypes> </ForceField>

#### □ Add machine/human readable SMARTS

o First atom token in SMARTS indicates the type that we are defining

Optionally, add colloquial description and DOI

o Adding DOI produces BibTeX file when applying force field





### **Promoting Dissemination of Force Fields**

- □ Template repo for building foyer compatible force field files
  - o https://github.com/mosdef-hub/forcefield\_template
  - o Skeleton force field file
  - o Skeleton testing setup
    - User needs to only add correctly typed .mol2 files
  - o Tutorial for adding SMARTS definitions
    - Walkthrough for adding more and more complex definitions
    - Testing with small example set of hydro carbons from alkanes through alkenes and benzene
    - How to run tests with py.test and enable TravisCl for your repo
- □ Force field development within our group
  - OPLS-AA compatible parameters for perfluoroethers

#### DOI 10.5281/zenodo.56807

- o Coarse-grained models for skin lipids
- o Automating force field derivation for supercapacitor electrodes



SCHOOL OF ENGINEERING





MUMS

### MetaMDS - Overview

□ MetaMDS was designed with the goal of being able to perform screening simulations

o It can submit large batches of jobs with various input parameters

#### □ MetaMDS relies on 3 key inputs:

- o mBuild "recipe"
  - This tells metaMDS how to construct the system
  - Allows us to screen over not just thermodynamic properties, but over different structures/molecules/etc.
- o Simulation template
  - This is a script that encapsulate the underlying simulation routines/calls
- o List(s) of parameters to vary
  - These parameters are used to change thermodynamic states in the simulation script and change inputs to the mBuild recipe
  - These lists could also include different forcefields to evaluate



### **MetaMDS - Screening**

- □ By using simulation templates, mBuild recipes, and simple parameter lists, we can create a streamlined, abstracted interface
  - o Users only need a few lines of code to do large scale, complex screening
  - o Individual elements can be easily tested before large scale deployment
  - o All procedures are well documented and described

```
In [ ]: # Initialize a simulation instance with a template and some metadata
                   sim = mds.Simulation(name='monolayer', template=create run script, output dir='output')
                   chain lengths = [8, 12, 16, 20]
                   for length in chain lengths:
                      parameters = { 'chain length': length,
                                    'n molecules': 100,
                                    'forcefield': 'OPLS-aa',
                                    'build func': build monolayer}
                       # Parameterize our simulation template
                       sim.parametrize(**parameters)
          In [ ]: # Run
                   sim.execute all(hostname='rahman.vuse.vanderbilt.edu', username='ctk3b')
          In [ ]: sim.sync all()
                                                                                                          CHOOL OF ENGINEERING
                              User interface is Python notebook (c.f. Mathematica)
```

#### Building a community

- o Hackathon in February with representation from 11 research groups and ExxonMobil
- o Goal of learning how to use MoSDeF and become co-developers
  - Add forcefields
    - Currently support OPLS, TRAPPE, specific forcefields (e.g., ionic liquids)
    - Wish to add other standard forcefields (e.g., bio forcefields, ReaxFF generic and specialized) plus workflows to automate forcefield derivation
  - Add simulation workflows for specific properties and applications
    - Currently support many standard workflows, coarse-graining (MSIBI), etc
    - LONG wish list!
  - Add simulation packages
    - Currently support LAMMPS, GROMACS, HOOMD-Blue
    - Wish to add open-source AIMD codes, Monte Carlo (e.g., Cassandra), other MD codes
- o Plan to jointly submit NSF SSI to support next MoSDeF cycle



SCHOOL OF ENGINEERING



# Putting it All Together: Signac Workflow Manager

Recently adopted in place of metaMDS

- Define heterogenous parameter spaces
- Provides link between data and metadata
  - o Every job and its data is tied to a point in parameter space
  - o Provides glue for seamless building, execution and analysis



Developed by Simon Adorf in Glotzer group

MUMS



SCHOOL OF ENGINEERING

Adorf et al., <u>arXiv:1611.03543</u> [cs.DB]



## **Putting It All Together**

#### Our tribology operations

- o mBuild + save with foyer
- o Minimize
- o Equilibrate
- o Shear at range of normal load

#### □ Customize compute environment

- o Defaults provided for common envs
- o Collection of envs for major computing centers
  - OLCF, NERSC, NICS

#### □ Execute and monitor



SCHOOL OF ENGINEERING

### Putting It All Together

#### Our tribology operations

- o mBuild + save with foyer
- o Minimize
- o Equilibrate
- o Shear at range of normal load

#### Customize compute environment

- o Defaults provided for common envs
- o Collection of envs for major computing centers
  - OLCF, NERSC, NICS

#### class MyTorqueEnvironment(flow.environment.TorqueEnvironment): hostname\_pattern = 'mymoabcluster.university.edu' cores\_per\_node = 16

#### @classmethod

def mpi\_cmd(cls, cmd, np):
 return 'mpirun -np {np} {cmd}'.format(n=np, cmd=cmd)

#### @classmethod

def script(cls, \_id, nn, walltime, ppn=None, \*\*kwargs):
 if ppn is None:
 ppn = cls.core\_per\_node
 js = super(MyTorqueEnvironment, cls).script()
 js.writeline('#/BS -i oe')
 js.writeline('#PBS -i nodes={}:ppn={}'.format(nn, ppn))
 js.writeline('#PBS -i walltime={}'.format(format\_timedelta(walltime)))
 js.writeline('#PBS -N {}'.format(\_id))
 js.writeline('#PBS -N {}'.format(\_id))
 return js

#### **Execute and monitor**



SCHOOL OF ENGINEERING

>>> project = MyProject() >>> project.print\_status(detailed=True, params=('a',)) Status project 'test-project': Total # of jobs: 10 label progress Detailed view: job\_id S next\_job a labels 108ef78ec381244447a108f931fe80db U 1 initialized, processed be01a9fd6b3044cf12c4a83ee9612f84 U 2 initialized, processed 32764c28ef130baefebeba76a158ac4e U process 3 initialized # ... >>>

### Intro to Nanotribology

- Micro- and nano-electromechanical devices (MEMS and NEMS) used in a wide range of applications
  - o Accelerometers
    - Airbag deployment
    - Consumer electronics
      - Apple iPhone, Nintendo Wii
  - o Micromirror arrays
    - HD projectors (DLPTV)
  - o Sensors
- Typically applications avoid parts in contact
  - o Surface forces such as adhesion, stiction, friction, and wear dominate
- Lubrication schemes needed for devices to reach full potential







\_

### **Monolayer Lubricants**

Considerations for nanoscale lubricants

- o Low surface energy
  - Minimize adhesion
- Low friction coefficient ( $\mu$ )
  - Minimize friction
- o Robust/durable
  - Wear-resistant
- □ Alkylsilane monolayers
  - o Densely-packed chains reduce surface energy, protect underlying surface from wear
    - Low adhesion between surfaces even in the presence of water
  - o Friction coefficients vary widely by composition, density, and chain length
    - $\mu$  reduced to < 50% of value for bare silica<sup>1</sup>
  - o Durability concerns



SCHOOL OF ENGINEERING

Alkanethiol monolayer on gold



Smith et al., Prog. in Surf. Sci., 2004

Alkylsilane monolayer on silica



Booth et al., Langmuir, 2011



### Screening Monolayers (with metaMDS)

#### Proposed hypotheses in the literature

- o More flexible chains reduce friction force on amorphous substrates
- o Crystalline monolayers imposed artificially induced ordering
  - Leads to lower friction coefficients

#### Design space:

- o Crystalline  $\beta$ -cristobalite and amorphous silica
- o Alkanes and polyethylene glycols (PEG)
- o Varied chain length and surface coverage







Klein, C., Sallai, J., **Jones, T. J**., lacovella, C. R., McCabe, C. Cummings, P.T., "A Hierarchical, Component Based Approach to Screening Properties of Soft Matter," In Foundations of Molecular Modeling and Simulation: Select Papers from FOMMS 2015; R. Q. Snurr; C. S. Adjiman and D.A. Kofke, Ed.; Springer Singapore: Singapore, 2016; pp 79-92.

### Nematic Order

- □ Nematic order measure of the global orientational ordering of monolayer chains
  - o 0 = isotropic
  - o I = perfectly ordered

$$S_2 = \langle \frac{3}{2} \cos^2 \theta - \frac{1}{2} \rangle$$

- PEG is found to feature lower nematic order than alkanes, particularly at lower surface coverage
  - o Floppier/more liquid-like





Klein, C., Sallai, J., **Jones, T. J.**, lacovella, C. R., McCabe, C. Cummings, P.T., "A Hierarchical, Component Based Approach to Screening Properties of Soft Matter," In Foundations of Molecular Modeling and Simulation: Select Papers from FOMMS 2015; R. Q. Snurr; C. S. Adjiman and D.A. Kofke, Ed.; Springer Singapore: Singapore, 2016; pp 79-92.





## Surface Coverage has Greater Impact on Short Chains

□ Subset of friction force data for polyethylene glycol monolayers on crystalline surface



Klein, C., Sallai, J., **Jones, T. J**., lacovella, C. R., McCabe, C. Cummings, P.T., "A Hierarchical, Component Based Approach to Screening Properties of Soft Matter," In Foundations of Molecular Modeling and Simulation: Select Papers from FOMMS 2015; R. Q. Snurr; C. S. Adjiman and D.A. Kofke, Ed.; Springer Singapore: Singapore, 2016; pp 79-92.

MUMS

SCHOOL OF ENGINEERING

### **Friction Forces**

 $\hfill\square$  Similar behavior for PEG and alkanes on amorphous surface

□ Strange results for alkanes on crystalline surfaces





### **Crystalline Monolayers Can Exhibit artifacts**

Crystalline monolayers impose artificial ordering on monolayers

- o Portions of monolayer can get trapped in different configurations
- o High forces perpendicular to shear direction

SCHOOL OF ENGINEERING





### **Terminal Group Chemistry**

- Terminal groups form the boundary at the monolayer interface, thus having important effects on tribology
  - o Methyl most common low surface energy
- □ Higher friction forces in hydroxyl-terminated monolayers
  - o Attributed to H-bonding across the monolayer interface
- Phenyl terminal groups can improve durability by resisting asperity penetration
- □ Only a few chemistries studied in the context of lubrication



Yu et al., Tribol. Lett., 2009

































 $F_{\text{friction}} = F_0 + \mu F_{\text{normal}}$ 





MUMS



42

### **Next Steps**

#### □ Large parameter set screening

- o Other factors which affect tribological behavior
  - Different tether chemistry
  - Mixed monolayers
    - Mixed length, end-functionalization, tether chemistry
  - Asymmetric functionalization different end groups on tethers
    - For N ways to produce a monolayer, N(N+1)/2 asymmetric functionalization scenarios
  - Identified ~45,000 variations to be screened
    - Hope to perform on OLCF Titan in the next few months

a. (





SCHOOL OF ENGINEERING

### Conclusions

- To perform MGI-style screening of soft materials requires automating many steps currently largely performed by hand
- □ MoSDeF is an environment for automating molecular simulations
  - o mBuild, Foyer, metaMDS,....
- □ Simulations performed within MoSDeF are completely reproducible
  - o All steps are scripted, which can be distributed online through github
    - In line with identified best practices for reproducibility of computational research
- MoSDeF is being co-developed with other groups interested in transparency, reproducibility, usability, and extensibility (TRUE)
- Application to nanoscale lubrication is one example of application of soft materials screening made possible by MoSDeF and other tools (Signac)
- $\Box$  All resources available on github with version control
  - o MoSDeF Hub https://github.com/mosdef-hub
  - o Foyer <u>https://github.com/mosdef-hub/foyer</u> and mBuild <u>https://github.com/mosdef-hub/mbuild</u>

SCHOOL OF ENGINEERING

### Acknowledgments

#### UVU ChBE

- o Clare McCabe
- Chris lacovella 0
- **Christoph Klein** 0
- Andrew Summers 0
- Justin Gilmer 0
- o Trevor Jones

#### 

- o Akos Ledezci
- Gabor Karsai 0
- Janos Sallai 0
- o Tengyu Ma

#### SCHOOL OF ENGINEERING

### Funding

- o NSF ACI 1047828: Development of an Integrated Molecular Design Environment for Lubrication Systems (iMoDELS)
- o NSF CDI CBET 1028374: Cyber-Enabled Design of **Functional Nanomaterials**
- o NSF SSE ACI 1535150: Development of a Software Framework for Formalizing Forcefield Atom-Typing for Molecular Simulation

#### □ Resources

- o National Energy Research Supercomputing Center (NERSC), Lawrence Berekeley Laboratory
- o Oak Ridge Leadership Class (OLC), Oak Ridge National Laboratory

### **MUM**