Incorporating the GAMESS Program Into the GAML Software

By: Santiago Cortes, Thomas Spangler.

Background

GAML was created by Professor Orlando Acevedo and his graduate student Xiang Zhong to automate the creation of force field (FF) parameters for use in molecular dynamics (MD) or Monte Carlo (MC) simulations

- The simulations are used to predict for any desired specific solvent molecule its precise macroscopic properties, such as:
  - density
  - heat of vaporization
  - free energy of hydration

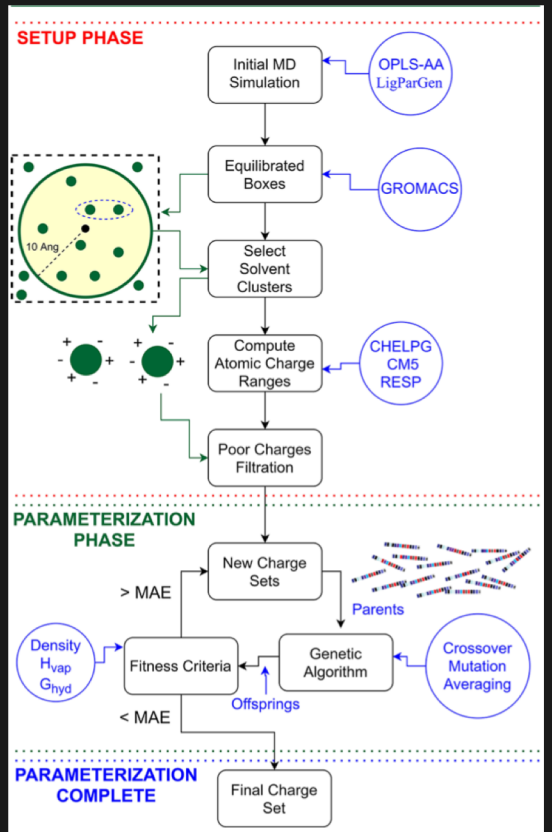


Figure 3. Schematic of the GAML software workflow.

GAML Explanation/Background

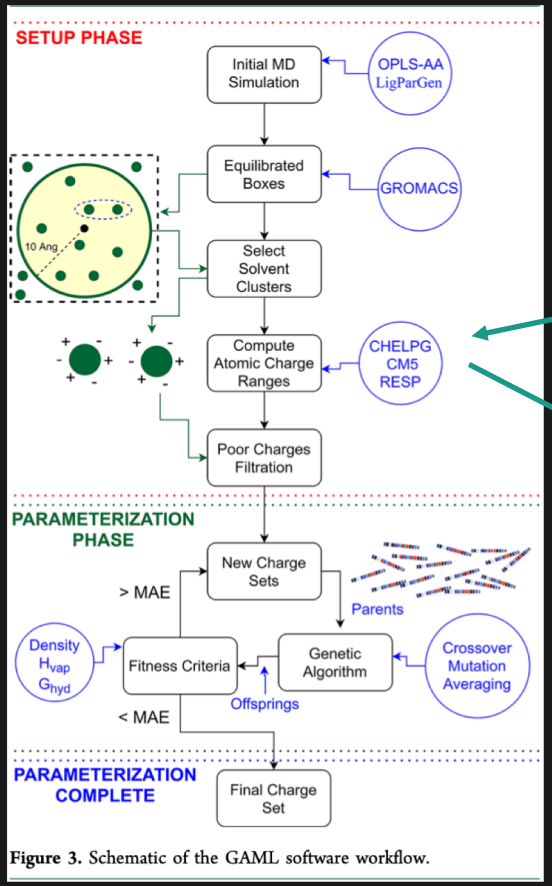


Figure 3. Schematic of the GAML software workflow.

$$E_{\text{bonds}} = \sum_i k_{r,i}(r_i - r_{0,i})^2 \quad (1)$$

$$E_{\text{angles}} = \sum_i k_{\theta,i}(\theta_i - \theta_{0,i})^2 \quad (2)$$

$$E_{\text{torsion}} = \frac{1}{2}\sum_i [V_{1,i}(1 + \cos\phi_i) + V_{2,i}(1 - \cos 2\phi_i) + V_{3,i}(1 + \cos 3\phi_i) + V_{4,i}(1 - \cos 4\phi_i)] \quad (3)$$

$$E_{\text{nonbond}} = \sum_i \sum_{j>i} \left\{ \frac{q_i q_j e^2}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]\right\} \quad (4)$$

```
REMARK LIGPARGEN GENERATED PDB FILE
ATOM      1  C00 MOL     1       1.000   1.000   0.000
ATOM      2  O01 MOL     1      -0.421   1.000   0.000
ATOM      3  C02 MOL     1      -0.942   1.000   1.322
ATOM      4  H03 MOL     1       1.344   1.001  -1.038
ATOM      5  H04 MOL     1       1.382   0.102   0.494
ATOM      6  H05 MOL     1       1.382   1.897   0.497
ATOM      7  H06 MOL     1      -2.033   0.999   1.262
ATOM      8  H07 MOL     1      -0.622   1.898   1.859
ATOM      9  H08 MOL     1      -0.619   0.103   1.860
END
```

```
 $SYSTEM TIMLIM=525600 $END
 $CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE COORD=PRINAXIS UNITS=ANGS $END
 $PDC PTSEL=CHELPG CONSTR=CHARGE $END
 $ELPOT IEPOT=1 WHERE=PDC $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
 $DATA
Dimethylether
C1
C  6.0  1.000  1.000  0.000
O  8.0  -0.421  1.000  0.000
C  6.0  -0.942  1.000  1.322
H  1.0  1.344  1.001  -1.038
H  1.0  1.382  0.102  0.494
H  1.0  1.382  1.897  0.497
H  1.0  -2.033  0.999  1.262
H  1.0  -0.622  1.898  1.859
H  1.0  -0.619  0.103  1.860
 $END
```

```
NET CHARGES:
-------------------------------------------
ATOM                CHARGE     E.S.D.
-------------------------------------------
C                    0.0518    0.6866
O                   -0.3892    0.0787
C                    0.0519    0.6866
H                    0.0631    0.0501
H                    0.0398    0.0492
H                    0.0398    0.0492
H                    0.0631    0.0501
H                    0.0398    0.0492
H                    0.0398    0.0492
-------------------------------------------
```

```
# Single-Point Charge Range for Dimethylether:

ATOM      1     0.0      0.3518
ATOM      2    -0.6892     -0.0892
ATOM      3     0.0      0.3519
ATOM      4     0.0      0.3631
ATOM      5     0.0      0.3398
ATOM      6     0.0      0.3398
ATOM      7     0.0      0.3631
ATOM      8     0.0      0.3398
ATOM      9     0.0      0.3398
```

Significance: Gaussian vs. GAMESS

$$\frac{-\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} = i\hbar\frac{\partial\Psi}{\partial t}$$

Schrodinger equation for free particle in one dimension.

Both:

- Compute partial atomic charges (atomic charge of an individual atom in a molecule)
- Require GPUs to operate (while working within GAML)

Gaussian:

- Already implemented into GAML
- Expensive commercial software package

GAMESS:

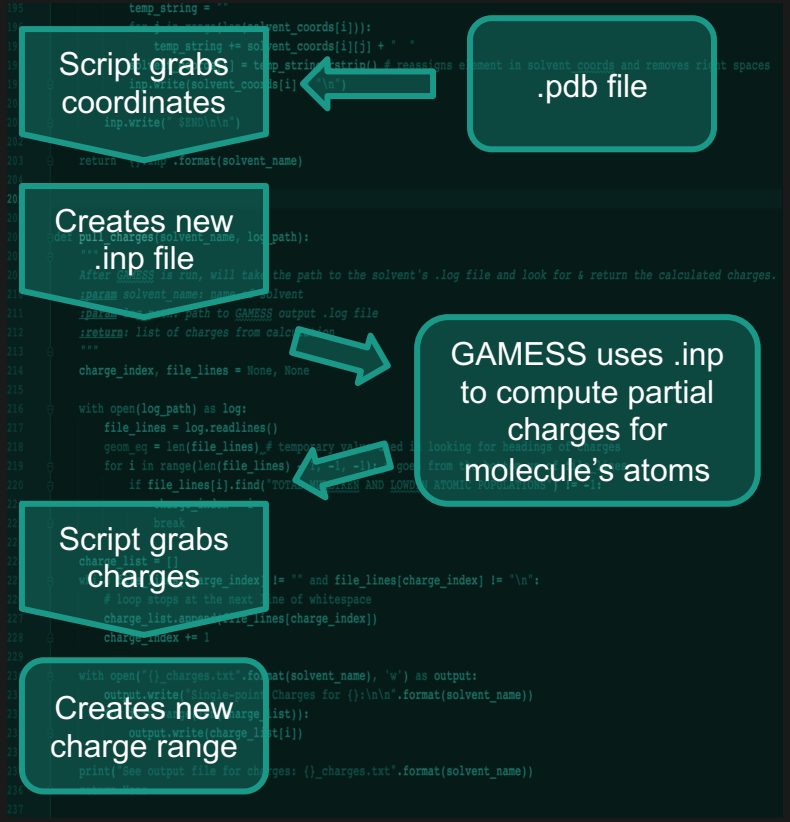
- Needs to be implemented
- Open-source

GAMESS

Project

We created and added new Python code to the GAML software to integrate GAMESS, which:

- locates a Gromacs file (.pdb) containing the user desired QM level of theory and the Cartesian coordinates that define the geometry of the solvent molecule
- extract that data to create an input file for GAMESS (.inp)
- Submits and completes the GAMESS calculation
- Parses the output file
- Returns a range of charges for the molecule, or CHELPG charges, needed for the GAML code



Outcomes

Our Python script was compatible with the GAML framework, in that it:

- located the specific information in the Gromacs input file
- created a new GAMESS file, submitted the GAMESS calculation
- located partial charges of the molecule
- updated GAML with the newly predicted atomic partial charges.

ex. - Dimethylether

```
REMARK LIGPARGEN GENERATED PDB FILE
ATOM      1  C00 MOL     1       1.000   1.000   0.000
ATOM      2  O01 MOL     1      -0.421   1.000   0.000
ATOM      3  C02 MOL     1      -0.942   1.000   1.322
ATOM      4  H03 MOL     1       1.344   1.001  -1.038
ATOM      5  H04 MOL     1       1.382   0.102   0.494
ATOM      6  H05 MOL     1       1.382   1.897   0.497
ATOM      7  H06 MOL     1      -2.033   0.999   1.262
ATOM      8  H07 MOL     1      -0.622   1.898   1.859
ATOM      9  H08 MOL     1      -0.619   0.103   1.860
END
```

```
 $SYSTEM TIMLIM=525600 $END
 $CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE COORD=PRINAXIS UNITS=ANGS $END
 $PDC PTSEL=CHELPG CONSTR=CHARGE $END
 $ELPOT IEPOT=1 WHERE=PDC $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
 $DATA
Dimethylether
C1
C  6.0  1.000  1.000  0.000
O  8.0  -0.421  1.000  0.000
C  6.0  -0.942  1.000  1.322
H  1.0  1.344  1.001  -1.038
H  1.0  1.382  0.102  0.494
H  1.0  1.382  1.897  0.497
H  1.0  -2.033  0.999  1.262
H  1.0  -0.622  1.898  1.859
H  1.0  -0.619  0.103  1.860
 $END
```

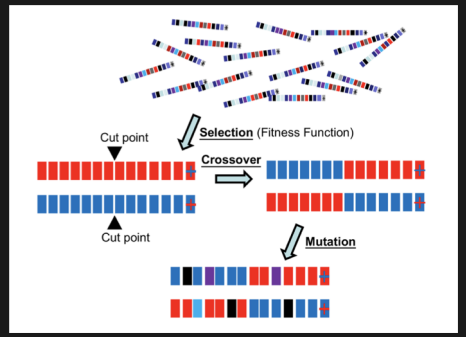
```
NET CHARGES:
-----------------------------------
ATOM                CHARGE    E.S.D.
-----------------------------------
C                   0.0518    0.6866
O                  -0.3892    0.0787
C                   0.0519    0.6866
H                   0.0631    0.0501
H                   0.0398    0.0492
H                   0.0398    0.0492
H                   0.0631    0.0501
H                   0.0398    0.0492
H                   0.0398    0.0492
-----------------------------------
```

```
# Single-Point Charge Range for Dimethylether:

ATOM      1     0.0     0.3518
ATOM      2     -0.6892     -0.0892
ATOM      3     0.0     0.3519
ATOM      4     0.0     0.3631
ATOM      5     0.0     0.3398
ATOM      6     0.0     0.3398
ATOM      7     0.0     0.3631
ATOM      8     0.0     0.3398
ATOM      9     0.0     0.3398
```

GAML Results

On the right is a table reflecting the results from GAML when ran with trial charge ranges calculated by GAMESS for chloroform, dimethylether, and pyridine; the trial values (red) were compared to the literature values (top black) and had error percentages calculated for each property on each solvent (bottom black).

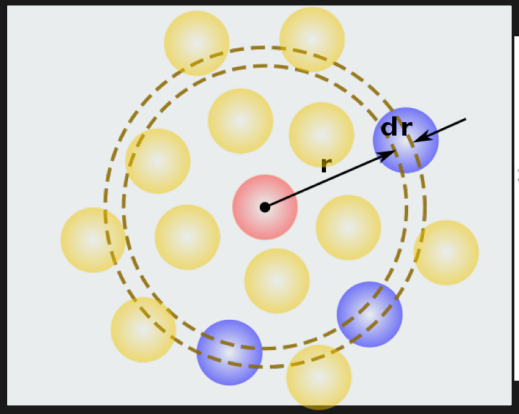


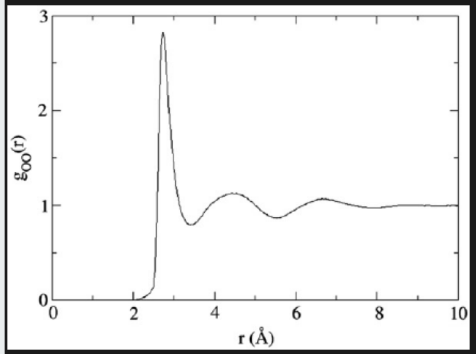
		Chloroform	Dimethylether	Pyridine
Literature Values	Density (kg/m^3)	1479	742	978
	Heats of Vaporization (kJ/mol)	31.28	21.72	40.15
	Free Energy of Hydration (kJ/mol)	-4.52	-7.99	-19.62
Simulation Results	Liquid	-9474.80	-5198.95	-2048.33
	Gas	10.028	11.455	38.524
	ΔH_{vap}	31.455	24.330	45.098
	Density	1513.71	727.38	1015.95
	ΔG_{hyd}	-4.250	-7.840	-18.021
	Box Size (nm)	4.03163	3.94953	4.01499
Percent Errors	Density	2.3%	-2.0%	3.9%
	ΔH_{vap}	0.6%	12.0%	12.3%
	ΔG_{hyd}	-6.0%	-1.9%	-8.2%
Mean Absolute Error		3.0%	5.3%	8.1%

Future work

For future work, the GAML program needs to:

- implement radial distribution functions (RDF) verification,
  - quantify the strength and structure of the intermolecular interactions present in the bulk phase
- be expanded to parameterize the nonbonded Lennard-Jones (LJ) terms of a FF
  - The LJ terms play a major role in RDF predictions.





$$E_{\text{bonds}} = \sum_i k_{r,i}(r_i - r_{0,i})^2 \quad (1)$$

$$E_{\text{angles}} = \sum_i k_{\theta,i}(\theta_i - \theta_{0,i})^2 \quad (2)$$

$$E_{\text{torsion}} = \frac{1}{2}\sum_i [V_{1,i}(1 + \cos\phi_i) + V_{2,i}(1 - \cos 2\phi_i) + V_{3,i}(1 + \cos 3\phi_i) + V_{4,i}(1 - \cos 4\phi_i)] \quad (3)$$

$$E_{\text{nonbond}} = \sum_i \sum_{j>i} \left\{ \frac{q_i q_j e^2}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \right\} \quad (4)$$

An additional direction would be to create a web-server version of GAML. This would expand accessibility to users lacking a GPU, as well as those with limited experience in Python, Linux, and command-line software.

Questions?

Trials

GAML Explanation/Background

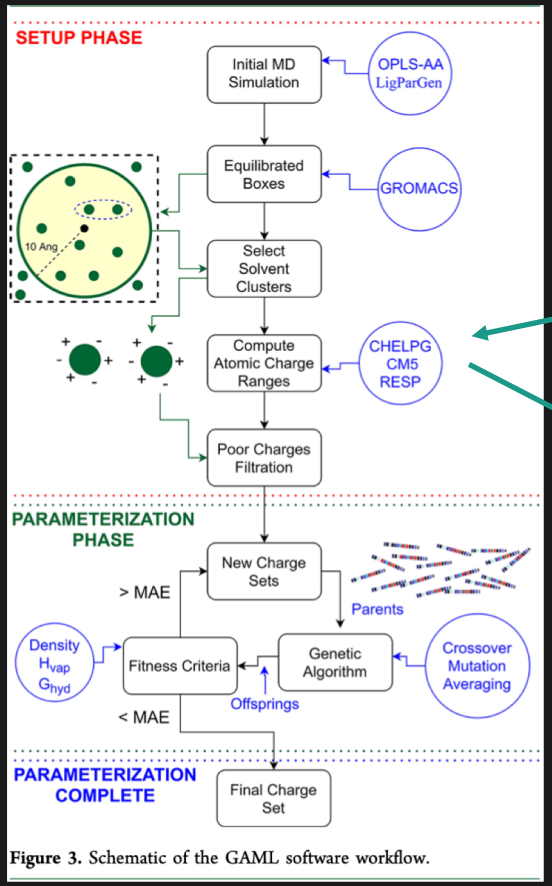


Figure 3. Schematic of the GAML software workflow.

$$E_{\text{bonds}} = \sum_i k_{r,i}(r_i - r_{0,i})^2 \quad (1)$$

$$E_{\text{angles}} = \sum_i k_{\theta,i}(\theta_i - \theta_{0,i})^2 \quad (2)$$

$$E_{\text{torsion}} = \frac{1}{2}\sum_i [V_{1,i}(1 + \cos\phi_i) + V_{2,i}(1 - \cos 2\phi_i) + V_{3,i}(1 + \cos 3\phi_i) + V_{4,i}(1 - \cos 4\phi_i)] \quad (3)$$

$$E_{\text{nonbond}} = \sum_i \sum_{j>i} \left\{ \frac{q_i q_j e^2}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]\right\} \quad (4)$$

```
REMARK LIGPARGEN GENERATED PDB FILE
ATOM      1  C00 MOL     1       1.000   1.000   0.000
ATOM      2  O01 MOL     1      -0.421   1.000   0.000
ATOM      3  C02 MOL     1      -0.942   1.000   1.322
ATOM      4  H03 MOL     1       1.344   1.001  -1.038
ATOM      5  H04 MOL     1       1.382   0.102   0.494
ATOM      6  H05 MOL     1       1.382   1.897   0.497
ATOM      7  H06 MOL     1      -2.033   0.999   1.262
ATOM      8  H07 MOL     1      -0.622   1.898   1.859
ATOM      9  H08 MOL     1      -0.619   0.103   1.860
END
```

```
 $SYSTEM TIMLIM=525600 $END
 $CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE COORD=PRINAXIS UNITS=ANGS $END
 $PDC PTSEL=CHELPG CONSTR=CHARGE $END
 $ELPOT IEPOT=1 WHERE=PDC $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
 $DATA
Dimethylether
C1
C  6.0  1.000  1.000  0.000
O  8.0  -0.421  1.000  0.000
C  6.0  -0.942  1.000  1.322
H  1.0  1.344  1.001  -1.038
H  1.0  1.382  0.102  0.494
H  1.0  1.382  1.897  0.497
H  1.0  -2.033  0.999  1.262
H  1.0  -0.622  1.898  1.859
H  1.0  -0.619  0.103  1.860
 $END
```

```
NET CHARGES:
-------------------------------------------
ATOM                CHARGE     E.S.D.
-------------------------------------------
C                    0.0518    0.6866
O                   -0.3892    0.0787
C                    0.0519    0.6866
H                    0.0631    0.0501
H                    0.0398    0.0492
H                    0.0398    0.0492
H                    0.0631    0.0501
H                    0.0398    0.0492
H                    0.0398    0.0492
-------------------------------------------
```

```
# Single-Point Charge Range for Dimethylether:

ATOM      1     0.0      0.3518
ATOM      2    -0.6892     -0.0892
ATOM      3     0.0      0.3519
ATOM      4     0.0      0.3631
ATOM      5     0.0      0.3398
ATOM      6     0.0      0.3398
ATOM      7     0.0      0.3631
ATOM      8     0.0      0.3398
ATOM      9     0.0      0.3398
```