QMC Workshop 2021

Observables
Week 6 \\ November 9th

Chandler Bennett, bennettcc@ornl.gov

https://github.com/QMCPACK/qmc_workshop_2021

Funding: U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials.

CPSFM Center for Predictive Simulation of Functional Materials

Sandia National Laboratories

Argonne NATIONAL LABORATORY

NC STATE UNIVERSITY

BROWN

OAK RIDGE National Laboratory

Outline

- Going beyond total energies
- The density estimator
  - VMC Case
  - DMC Case – Mixed Estimator Bias
- Overview of lab material (Density of Water)

Next Week: QMC research presentations from guest lecturers

Going Beyond Total Energies

Other Scalar Quantities (QMCA Tool)

- The ***scalar.dat*** files of QMCPACK record numerous quantities per block (not just total energy and variance) :

```
$ head vmc.s000.scalar.dat
#   index    LocalEnergy         LocalEnergy_sq      LocalPotential      Kinetic             ElecElec            IonIon              LocalECP            NonLocalECP         BlockWeight         BlockCPU            AcceptRatio
        0   -1.0391434585e+01    1.0950516390e+02   -2.1419685812e+01    1.1028251227e+01   -2.5317857117e+00   -1.2775667432e+01   -6.7934302226e+00    6.8119755391e-01    3.0000000000e+03    7.3457234502e+00    9.2262500000e-01
        1   -1.0411339626e+01    1.0988693409e+02   -2.1774462570e+01    1.1363122945e+01   -2.5034713608e+00   -1.2775667432e+01   -7.2837665515e+00    7.8844277359e-01    3.0000000000e+03    7.3943826358e+00    9.2083333333e-01
        2   -1.0379473128e+01    1.0909336095e+02   -2.1330053182e+01    1.0950580054e+01   -2.5773455834e+00   -1.2775667432e+01   -6.8560422936e+00    8.7900212655e-01    3.0000000000e+03    7.3868751327e+00    9.2437500000e-01
        3   -1.0358077902e+01    1.0905398586e+02   -2.1676398015e+01    1.1318320113e+01   -2.4642489678e+00   -1.2775667432e+01   -7.2479079313e+00    8.1142631523e-01    3.0000000000e+03    7.3507052263e+00    9.2045833333e-01
        4   -1.0399365618e+01    1.0972530108e+02   -2.1760641971e+01    1.1361276353e+01   -2.4806183383e+00   -1.2775667432e+01   -7.4264592338e+00    9.2210303217e-01    3.0000000000e+03    7.4179414709e+00    9.2441666667e-01
        5   -1.0404822949e+01    1.0959064697e+02   -2.1899338555e+01    1.1494515605e+01   -2.4796978861e+00   -1.2775667432e+01   -7.4662338348e+00    8.2226059785e-01    3.0000000000e+03    7.4227228562e+00    9.1820833333e-01
        6   -1.0314305494e+01    1.0842636255e+02   -2.1540517834e+01    1.1226212340e+01   -2.3982356455e+00   -1.2775667432e+01   -7.1506077195e+00    7.8399296220e-01    3.0000000000e+03    7.4203772346e+00    9.1962500000e-01
        7   -1.0389331065e+01    1.0951191701e+02   -2.1671312218e+01    1.1281981153e+01   -2.5114567027e+00   -1.2775667432e+01   -7.1806554097e+00    7.9646732568e-01    3.0000000000e+03    7.3896697164e+00    9.2087500000e-01
        8   -1.0442567254e+01    1.1050816567e+02   -2.1255908027e+01    1.0813340774e+01   -2.6784232830e+00   -1.2775667432e+01   -6.4695998449e+00    6.6778253198e-01    3.0000000000e+03    7.2716093063e+00    9.2675000000e-01
```

- The means and error bars can be obtained with the ***qmca*** tool

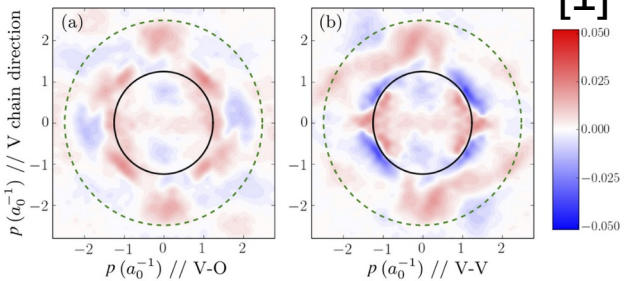
```
$ qmca vmc.s000.scalar.dat

vmc  series 0
  LocalEnergy           =          -10.3968 +/-           0.0025
  Variance              =             1.666 +/-           0.021
  Kinetic               =            11.206 +/-           0.011
  LocalPotential        =           -21.603 +/-           0.011
  ElecElec              =           -2.5169 +/-          0.0039
  LocalECP              =            -7.135 +/-           0.015
  NonLocalECP           =            0.8249 +/-          0.0050
  IonIon                =            -12.78 +/-            0.00
  LocalEnergy_sq        =           109.761 +/-           0.050
  BlockWeight           =           3000.00 +/-            0.00
  BlockCPU              =            7.4013 +/-          0.0035
  AcceptRatio           =           0.92320 +/-         0.00013
  Efficiency            =            128.81 +/-            0.00
  TotalTime             =           1480.26 +/-            0.00
  TotalSamples          =            600000 +/-               0
```

Beyond Energies

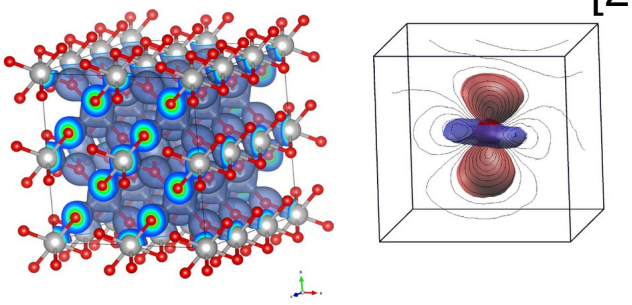
- QMCPACK is capable of calculating a broad range of observables (estimators), not just the total energy (e.g. the Chiesa KE correction shown last week).
- Others include:
  - **Energy Density**
  - **One-body reduced Density Matrix**
  - **Spherically Averaged Pair-Correlation**
  - **Static Structure Factor**
  - **Forces**
  - **And More …**
- For a detailed list and their formal descriptions and implementation details, see the following:
  - https://qmcpack.readthedocs.io (Online Documentation)
  - https://doi.org/10.1088/1361-648X/aab9c3 (Citation Paper)
  - https://doi.org/10.1063/5.0004860 (QMCPACK Advances)

[1]



Momentum Distribution

[2,3]



Density and Spin-Density

[1] Phys. Rev. B 99, 075154 (2019)
[2] New J. Phys. 18 113049 (2016)
[3] cond-mat arXiv:2103.09809 (2021)

Notes on QMCPACK Estimators

- To calculate estimators, modifications to the QMCPACK input are generally minor, intuitive, and user friendly:

```
<estimator name="Density" type="density" delta="0.05 0.05 0.05"/>  <!-- Density Estimator -->
<estimator type="spindensity" name="SpinDensity" report="yes">     <!-- Spin-Density Estimator -->
    <parameter name="grid"> 40 40 40 </parameter>
</estimator>
<estimator type="gofr" name="gofr" num_bin="200" rmax="3.0" />     <!-- Pair-Correlation Estimator -->
<estimator type="sk" name="sk" hdf5="yes"/>                        <!-- Static Structure Factor Estimator -->
```

- Estimators can be evaluated during all chained QMC runs, for instance: **VMC ➡ DMC ➡ DMC ➡ ...**
- However, the estimator elements can also be evaluated at specific QMC sections (i.e. for VMC only). The online documentation outlines precisely how this is specified.
- Generally not available (or needed) for optimizations.

The Density Estimator

The Density Estimator – VMC Case

- To calculate the density, the particle number density operator is used. And its form is given as:

$$\hat{n}_{\vec{r}} = \sum_{i=1}^{N} \delta(\vec{r} - \vec{r}_i), \quad \textbf{Particle Number Density Operator}$$

- Provides a count of the number of electrons at a point $\vec{r}$.
- In practice, QMCPACK must use a finite histogram grid, then average the number of particles in each grid cell over an ensemble of walkers distributed as $\Psi_T(\vec{R})$ to obtain the density $n(\vec{r})$:

$$n_c = \int \mathrm{d}^3 R |\Psi_T(\vec{R})|^2 \int_{\Omega_c} \mathrm{d}^3 r \sum_{i=1}^{N} \delta(\vec{r} - \vec{r}_i)$$
$$\approx \frac{1}{M} \sum_{m=1}^{M} [\#\text{ of sample's electrons in cell, } c]$$

Note:
The **spin-density** is obtained similarly; up and down electrons are counted *separately* to generate up and down densities.

The Density Estimator – DMC Case

- First, let's briefly step back and look closely at how the energy is evaluated in DMC:
- **Mixed-Estimator:**

$$E = \lim_{\tau\to\infty} \frac{< e^{-\tau\hat{H}/2}\Psi_T|\hat{H}|e^{-\tau\hat{H}/2}\Psi_T >}{< e^{-\tau\hat{H}/2}\Psi_T|e^{-\tau\hat{H}/2}\Psi_T >}$$

$$= \lim_{\tau\to\infty} \frac{< e^{-\tau\hat{H}}\Psi_T|\hat{H}|\Psi_T >}{< e^{-\tau\hat{H}}\Psi_T|\Psi_T >} = \frac{< \Psi_0|\hat{H}|\Psi_T >}{< \Psi_0|\Psi_T >}$$

$$= \frac{\int \mathrm{d}^3R\, \Psi_0(\vec{R})\Psi_T(\vec{R})E_L(\vec{R})}{\int \mathrm{d}^3R\, \Psi_0(\vec{R})\Psi_T(\vec{R})}$$

(1) Use commuting property

(2) "Mixed Estimator"

(3) Average of local energy over dist. $\Psi_0\Psi_T$.

- Doesn't work for observables that do not commute with H

The Density Estimator – DMC Case

- If we evaluate an observable, O, that does not commute with H in the same way as the energy, an error is incurred $\mathcal{O}[(\Psi_0 - \Psi_T)]$.
- For this reason, in practice a combination of the mixed estimator (DMC) and the variational estimator (VMC) is used:

$$< \Psi_0|\mathcal{O}|\Psi_0 > \approx 2 < \Psi_0|\mathcal{O}|\Psi_T > - < \Psi_T|\mathcal{O}|\Psi_T >$$

(DMC) **(VMC)**

(Extrapolated Estimator)

- Notice that the accuracy of the extrapolated estimator depends on the quality of the trial wave function. Its error now $\mathcal{O}[(\Psi_0 - \Psi_T)^2]$.
- An alternate approach to get around this problem is to use a method known as “forward walking” which enables one to sample the pure distribution $< \Psi_0|\Psi_0 >$. This is beyond the scope of this talk, though the method is implemented in QMCPACK and details can be found in the online manual.

Notes:

- When calculating the density for systems with open boundaries, the user must specify the cell that will be used to evaluate the density – best to center the molecule in the desired cell. For periodic boundaries, the supercell is used by default.
- Be sure that when estimating any observable, if it does not commute with H then you should use an extrapolated estimate (KE, Potential Energy, etc.)

Questions?

Overview of Lab Material:

Calculating and analyzing the density of water

Hands-on Tutorial

Step 0: **git pull** inside **qmc_workshop_2021** and **~/apps/qmcpack/qmcpack**

Step 1: Orbital and wave function generation using PySCF and QMCPACK

Step 2: Calculating and storing VMC and DMC charge density

Step 3: Using **qdens** tool to generate .xsf file for visualization in VESTA

Step 4: Using **qdens-radial** tool to estimate atomic occupations (with an extrapolated estimate and resampled error bars)

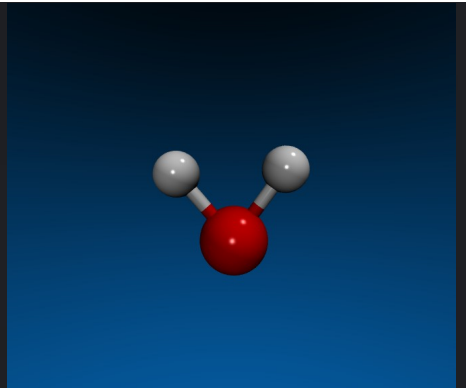
Hands-on Tutorial

Step 1: Orbital and wave function generation using PySCF and QMCPACK

- Material located at: *./qmc_workshop_2021/week6_observables/01_density_water*
- Nexus work flow file: *h2o_workflow.py*
- Define the pseudopotentials and structure

```
ppset(
    label    = 'ccecp',
    qmcpack = ['H.ccECP.xml','O.ccECP.xml'],
    )

system = generate_physical_system(
    structure = 'H2O.xyz',  # H2O atomic structure
    H         = 1,          # H pseudo Zeff
    O         = 6,          # O pseudo Zeff
    )
```



Hands-on Tutorial

Step 1: Orbital and wave function generation using PySCF and QMCPACK

- Hartree-Fock (HF) and Orbital Conversion

```
# perform Hartree-Fock
scf = generate_pyscf(
    identifier = 'scf',              # log output goes to scf.out
    path       = 'H2O/hf',           # directory to run in
    job        = job(serial=True,app='python3'),
    template   = './scf_template.py', # pyscf template file
    system     = system,
    mole       = obj(                # used to make Mole() inputs
        ecp      = 'ccecp',
        basis    = 'ccecp-ccpvqz',
        symmetry = True,
        ),
    save_qmc   = True,               # save wfn data for qmcpack
    )

# convert orbitals to QMCPACK format
c4q = generate_convert4qmc(
    identifier   = 'c4q',
    path         = 'H2O/hf',
    job          = job(cores=1),
    no_jastrow   = True,
    dependencies = (scf,'orbitals'),
    )
```

- converged SCF energy = -16.9177692534032

Hands-on Tutorial

Step 1: Orbital and wave function generation using PySCF and QMCPACK

- Two-body Jastrow optimization

```
# optimize 2-body Jastrow
optJ2 = generate_qmcpack(
    identifier      = 'opt',
    path            = 'H2O/optJ2',
    job             = job(cores=cores),
    system          = system,
    pseudos         = 'ccecp',
    J2              = True,
    J2_rcut         = 8.0,
    qmc             = 'opt',          # use opt defaults
    minmethod       = 'oneshiftonly', # adjust for oneshift
    init_cycles     = 3,
    init_minwalkers = 0.1,
    cycles          = 3,
    samples         = 25600,
    dependencies    = [(c4q,'particles'),(c4q,'orbitals' )],
    )
```

- In this case, HF leads to a wave function with a variance-to-energy ratio of ~0.02 Ha.

Hands-on Tutorial

Step 2: Calculating and storing VMC and DMC charge density

- Defining the QMC runs with density estimator:

```
from qmcpack_input import density
# run DMC with Slater-Jastrow wavefunction
qmc = generate_qmcpack(
    identifier   = 'dmc',
    path         = 'H2O/dmc_J2',
    job          = job(cores=cores),
    system       = system,
    pseudos      = 'ccecp',
    jastrows     = [],
    estimators = [density(delta=(0.02,0.02,0.02),x_min=0,x_max=8,y_min=0,y_max=8,z_min=0,z_max=8)], # Grid spacing in bohr
    seed         = 42,
    qmc          = 'dmc',     # DMC run
    vmc_samples   = 1024,     # DMC walker population sampled from VMC
    vmc_blocks    = 150,
    vmc_steps     = 25,
    vmc_timestep  = 0.3,
    eq_dmc        = True,     # Add DMC equilibration
    eq_blocks     = 30,       # Use a small number of blocks
    eq_steps      = 10,
    eq_timestep   = 0.02,     # Use a larger timestep
    blocks        = 500,      # Large number of blocks for production
    steps         = 25,       # 10 steps/block averages out some autocorr time
    timestep      = 0.01,     # Smaller production timestep
    nonlocalmoves = True,     # Use T-moves scheme w/ non-local pseudopotentials
    dependencies = orbdeps+[(optJ2,'jastrow')],
    )
```

Hands-on Tutorial

Step 3: Using **qdens** tool to generate .xsf file for visualization in VESTA

- After the QMC runs complete, the density for each block will be located in the *dmc.s*.stat.h5* files.
- We use the **qdens** for two reasons: (1) average density over all the blocks and store the average density in a format suitable for visualization and/or analysis.
- **Note:** Use **qmca** first to determine equilibration and auto-correlation lengths

```
# VMC (Series 000)
qdens -f xsf -e 10 -r 1 -i dmc.in.xml --noplot dmc.s000.stat.h5
```

File Format | # of Equil. Blocks | Reblock Factor | QMCPACK Input File | No plotting | .h5 file containing Density

Command will generate 3 files:

1. dmc.s000.Density_q.xsf **(Mean)**
2. dmc.s000.Density_q+err.xsf
3. dmc.s000.Density_q-err.xsf

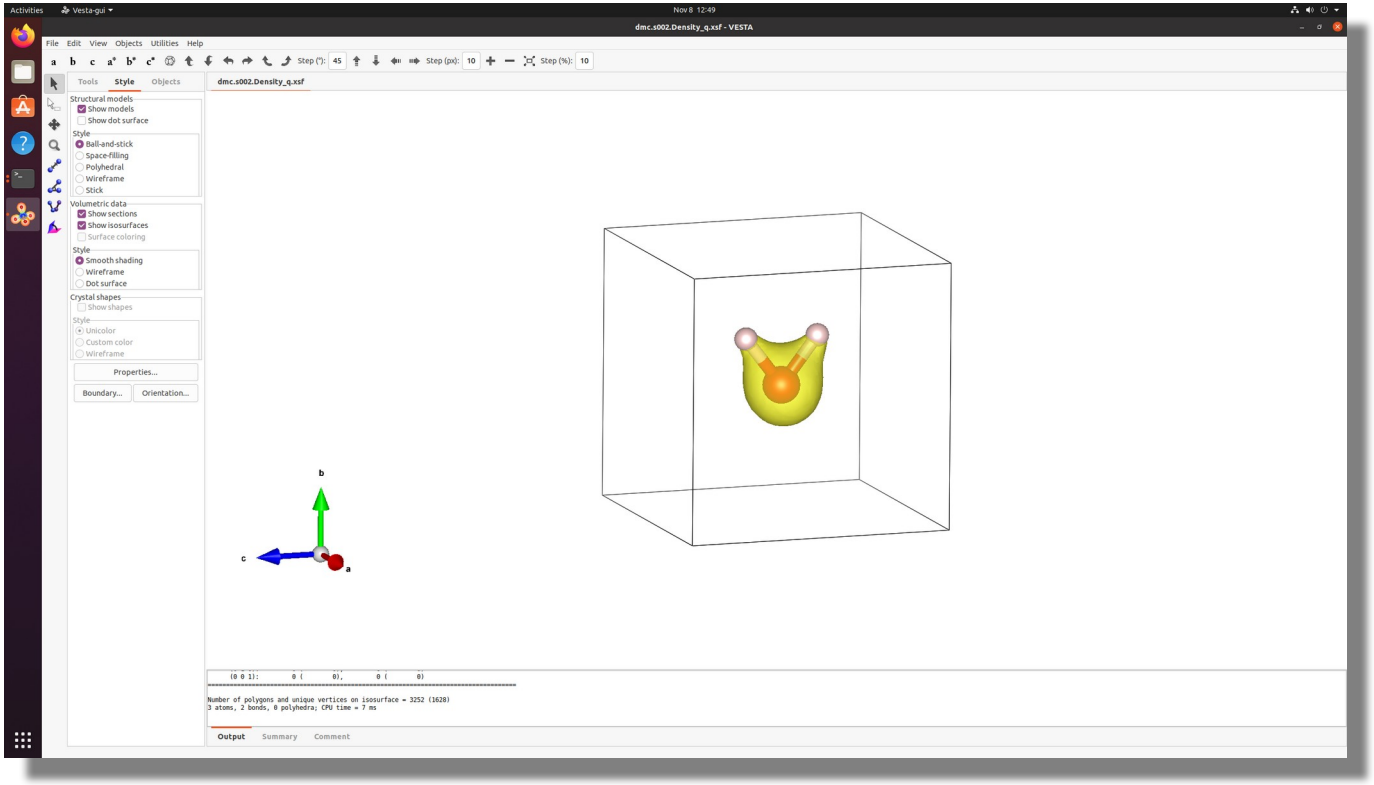
```
# DMC (Series 002)
qdens -f xsf -e 25 -r 3 -i dmc.in.xml --noplot dmc.s002.stat.h5
```

Command will generate 3 files:

1. dmc.s002.Density_q.xsf **(Mean)**
2. dmc.s002.Density_q+err.xsf
3. dmc.s002.Density_q-err.xsf

Hands-on Tutorial

Step 3: Using **qdens** tool to generate .xsf file for visualization in VESTA



To visualize:

$ VESTA

Then import .xsf file:

File > Open > dmc.s002.Density_q.xsf

VESTA Manual:

https://jp-minerals.org/vesta/en/doc.html

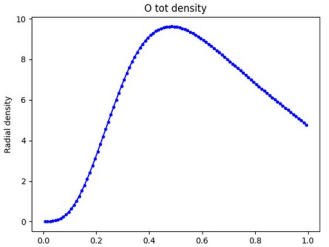
Hands-on Tutorial

Step 4: Using **qdens-radial** tool to estimate atomic occupations

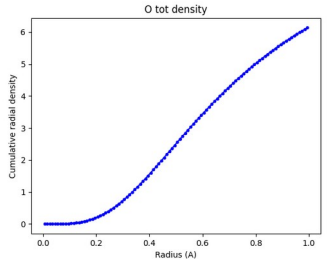
- Once we have the *Density*.xsf files from qdens, we can use another QMCPACK tool, **qdens-radial**, to calculate on-site populations.
- How does this work: given a set of species and radii (in Angstroms), the tool will generate a radial density – angular average – around the atomic sites out to the specified radius
- The radial density can be chosen to be non-cumulative or cumulative (integrated).

 # Example 1: Plot DMC non-cumulative radial density of O.
qdens-radial -p -s O -r 1 dmc.s002.Density_q.xsf

Plot Species Radii (A) .xsf file containing Density



 # Example 2: Plot DMC cumulative radial density of O.
qdens-radial -p -s O -r 1 **-c** dmc.s002.Density_q.xsf



Hands-on Tutorial

Step 4: Using **qdens-radial** tool to estimate atomic occupations

- For the cumulative case, **qdens-radial** will also print the cumulative value at the specified radius, i.e., an estimate of the atomic occupation

```
# Example 3: Estimate of the DMC atomic occupation
qdens-radial -p -s O -r 1.1 -c dmc.s002.Density_q.xsf
```

Output:

```
Cumulative Value of O Species at Cutoff 1.0 is: 6.55517033828574
```

- We can also get an extrapolated estimate (mixed-estimator bias) for this quantity by providing a VMC .xsf file:

```
# Example 4: Estimate of the extrapolated atomic occupation
qdens-radial -p -s O -r 1.1 -c --vmc=dmc.s000.Density_q.xsf dmc.s002.Density_q.xsf
```

Output:

```
Extrapolating from VMC and DMC densities...
Cumulative Value of O Species at Cutoff 1.1 is: 6.576918233167152
```

Hands-on Tutorial

Step 4: Using **qdens-radial** tool to estimate atomic occupations

- What about error bars?
- We can “resample” the density at each grid point:
- Recipe:
  1. Use error bars from ***.Density_q+err.xsf** file, draw samples from a Gaussian distribution with standard deviation that matches error bar.
  2. Calculate occupations with resampled data and calculate standard deviation to obtain the error bar on the occupation
  3. Make sure number of samples is converged

 # Example 5: Estimate DMC atomic occupation with error bar

qdens-radial -p -s O -r 1.1 -c **-n 20 --dmcerr=dmc.s002.Density_q+err.xsf** dmc.s002.Density_q.xsf

Output:

Resampling to obtain error bar (NOTE: This can be slow)...
Will compute 20 samples...
...
Cumulative Value of O Species at Cutoff 1.1 is: **6.55517033828574+/-0.0015585537493962 79**

Notes:

Questions?