

Monte Carlo Statistical Analysis: Averages, Error Bars, Variance

William Parker (ANL)

Background

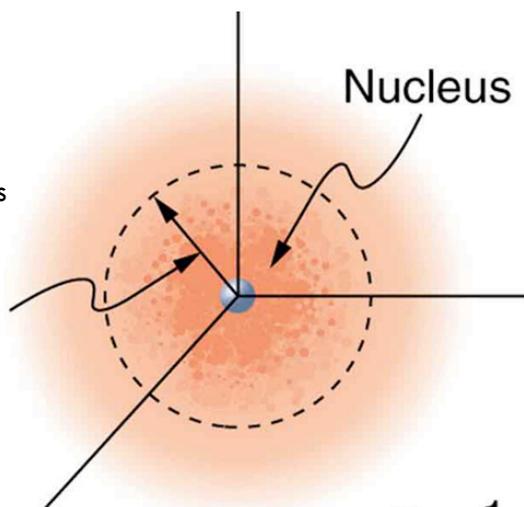
Atomic units:

Length : 1 bohr = 0.529 Å
= 52.9 pm



N. Bohr –
Wikimedia Commons

$r_1 = a_B$
Most probable
distance for
the electron



$$\begin{aligned} n &= 1 \\ l &= 0 \\ m_l &= 0 \end{aligned}$$

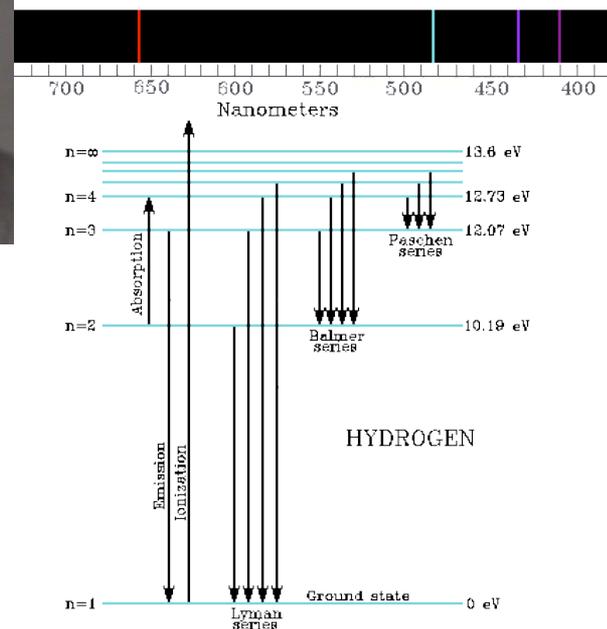
<http://cnx.org/content/m42606/latest/>

Charge: $1e = 1.61 \times 10^{-19} \text{ C}$
Action: $1\hbar = 1.05 \times 10^{-34} \text{ J s}$

Energy : 1 hartree = 27.2 eV
= 4.36 aJ



D. R. Hartree –
National Portrait
Gallery (mw221452)



<http://www.physics.udel.edu/~watson/scen103/colloq2000/images/hydrogen.gif>

Mass: $1 m_e = 9.11 \times 10^{-31} \text{ kg}$
Time: $1\hbar/\text{hartree} = 24.1 \text{ as}$

Background

Ground state of H atom

(in atomic units)

$$\hat{H} = \hat{T} + \hat{V}$$
$$\left(-\frac{1}{2} \nabla^2 - \frac{1}{r} \right) \psi = E \psi$$

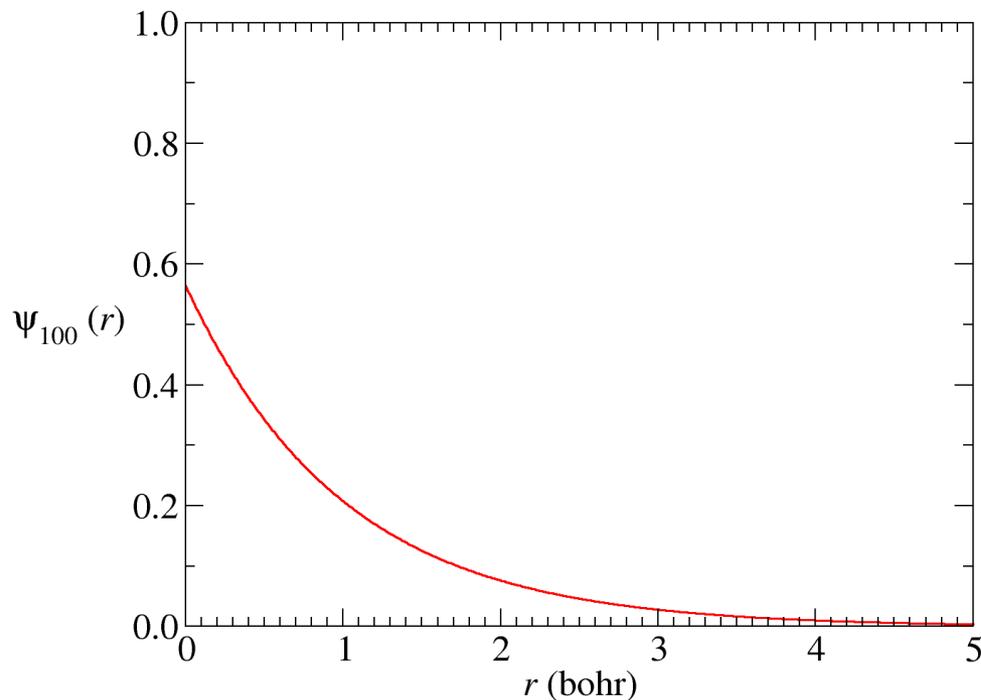
$$\psi_{100}(r) = \frac{1}{\sqrt{\pi}} e^{-r}$$

$$E_1 = -0.5 \text{ hartree}$$

$$\langle \hat{T} \rangle = 0.5 \text{ hartree}$$

$$\langle \hat{V} \rangle = -1.0 \text{ hartree}$$

$$\langle \hat{V} \rangle = -2 \langle \hat{T} \rangle$$



$$E_L = \frac{\hat{H} \psi}{\psi}$$

Outline

1. Reviewing how to calculate average and standard deviation
2. Looking at MC data in **QMCPACK**'s scalar.dat file
3. Averaging MC quantities using qmca
4. Evaluating MC simulation quality by tracing, autocorrelation & variance
5. Reducing error bars on MC averages with increased sampling and improved wave functions
6. Calculating MC simulation efficiency and scaling with particle number

Reviewing average and standard deviation

Average (mean): (of N samples)

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Standard deviation (error on known mean): (of N samples)

$$\sigma = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Standard deviation (unknown mean): (of N samples)

Use second moment (x^2) to estimate

$$\tilde{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [(x^2)_i - (x_i)^2]}$$

Looking at MC data

QMCPACK outputs data to (prefix).scalar.dat file [also: (prefix).dmc.dat for DMC simulations]

$$i \quad \left(\frac{\hat{H}\psi}{\psi} \right)_i \quad \left(\frac{\hat{H}^2\psi}{\psi} \right)_i \quad \left(\frac{\hat{V}\psi}{\psi} \right)_i \quad \left(\frac{-\frac{1}{2}\nabla^2\psi}{\psi} \right)_i \quad \left(-\sum_{R_{e-ion}} \frac{1}{R_{e-ion}} + \sum_{r_{e-e}} \frac{1}{r_{e-e}} \right)_i \quad w_i \quad (t_{\text{CPU}})_i \quad \left(\frac{N_{\text{accept}}}{N_{\text{propose}}} \right)_i$$

index	LocalEnergy	LocalEnergy_sq	LocalPotential	Kinetic	Coulomb	BlockWeight	BlockCPU	AcceptRatio
0	-4.5298911858e-01	6.0984565298e-01	-1.1708693521e+00	7.1788023352e-01	-1.1708693521e+00	1.2800000000e+04	6.0178991748e-03	9.8515625000e-01
1	-4.5481953564e-01	6.1641291630e-01	-1.1863425644e+00	7.3152302871e-01	-1.1863425644e+00	1.2800000000e+04	5.8323097461e-03	9.8562500000e-01
2	-4.8066105923e-01	1.3586015116e+00	-1.1766446209e+00	6.9598356165e-01	-1.1766446209e+00	1.2800000000e+04	5.8213412744e-03	9.8531250000e-01
3	-4.7316713469e-01	7.8720769003e-01	-1.1799481122e+00	7.0678097751e-01	-1.1799481122e+00	1.2800000000e+04	5.8330412549e-03	9.8828125000e-01
4	-4.6204733302e-01	5.6393677687e-01	-1.1619244081e+00	6.9987707508e-01	-1.1619244081e+00	1.2800000000e+04	5.8108362256e-03	9.8625000000e-01
5	-4.4313854290e-01	6.0831516179e-01	-1.2064503041e+00	7.6331176120e-01	-1.2064503041e+00	1.2800000000e+04	5.8254170264e-03	9.8625000000e-01
6	-4.5064926960e-01	5.9891422196e-01	-1.1521370176e+00	7.0148774798e-01	-1.1521370176e+00	1.2800000000e+04	5.8314813086e-03	9.8679687500e-01
7	-4.5687452611e-01	5.8139614676e-01	-1.1423627617e+00	6.8548823555e-01	-1.1423627617e+00	1.2800000000e+04	5.8258469971e-03	9.8726562500e-01
8	-4.5018503739e-01	8.4147849706e-01	-1.1842075439e+00	7.3402250655e-01	-1.1842075439e+00	1.2800000000e+04	5.8158433545e-03	9.8468750000e-01
9	-4.3862013841e-01	5.5477715836e-01	-1.2080979177e+00	7.6947777925e-01	-1.2080979177e+00	1.2800000000e+04	5.7959401123e-03	9.8539062500e-01
10	-4.5791153007e-01	6.0138927970e-01	-1.1353088208e+00	6.7739729068e-01	-1.1353088208e+00	1.2800000000e+04	5.8204662207e-03	9.8562500000e-01
11	-4.6353069662e-01	8.1222513185e-01	-1.1816447482e+00	7.1811405154e-01	-1.1816447482e+00	1.2800000000e+04	5.8378427930e-03	9.8664062500e-01
12	-4.5865881127e-01	7.0268353090e-01	-1.1601220665e+00	7.0146325525e-01	-1.1601220665e+00	1.2800000000e+04	5.8299140234e-03	9.8656250000e-01
13	-4.3846302063e-01	7.4457683787e-01	-1.2563266080e+00	8.1786358741e-01	-1.2563266080e+00	1.2800000000e+04	5.8239151855e-03	9.8632812500e-01
14	-4.5068572293e-01	6.0226458066e-01	-1.1888498591e+00	7.3816413622e-01	-1.1888498591e+00	1.2800000000e+04	5.8064188281e-03	9.8539062500e-01
15	-4.3816209316e-01	5.6515677610e-01	-1.1499660387e+00	7.1180394558e-01	-1.1499660387e+00	1.2800000000e+04	5.8297365283e-03	9.8617187500e-01
16	-4.7241181255e-01	7.0149650305e-01	-1.1523315578e+00	6.7991974526e-01	-1.1523315578e+00	1.2800000000e+04	5.8177384424e-03	9.8531250000e-01
17	-4.5302636441e-01	5.5508361049e-01	-1.1426585165e+00	6.8963215208e-01	-1.1426585165e+00	1.2800000000e+04	5.8202933984e-03	9.8460937500e-01
18	-4.5782213514e-01	7.6487410290e-01	-1.2030357453e+00	7.4521361013e-01	-1.2030357453e+00	1.2800000000e+04	5.8206455615e-03	9.8515625000e-01
19	-4.6635037994e-01	1.8182574986e+00	-1.1761497455e+00	7.0979936558e-01	-1.1761497455e+00	1.2800000000e+04	5.8067459180e-03	9.8640625000e-01
20	-4.5841707138e-01	7.0825060159e-01	-1.2012275308e+00	7.4281045937e-01	-1.2012275308e+00	1.2800000000e+04	5.8285895361e-03	9.8656250000e-01
21	-4.5936382993e-01	6.8521807935e-01	-1.2017023426e+00	7.4233851270e-01	-1.2017023426e+00	1.2800000000e+04	5.8254068115e-03	9.8570312500e-01
22	-4.4023825329e-01	5.9781382546e-01	-1.2094503406e+00	7.6921208728e-01	-1.2094503406e+00	1.2800000000e+04	5.8299661035e-03	9.8710937500e-01
23	-4.5411428000e-01	4.9904220387e-01	-1.1350953586e+00	6.8098107860e-01	-1.1350953586e+00	1.2800000000e+04	5.8216890771e-03	9.8523437500e-01
24	-4.6486624346e-01	5.3761062470e-01	-1.1095842543e+00	6.4471801080e-01	-1.1095842543e+00	1.2800000000e+04	5.8031559424e-03	9.8515625000e-01
25	-4.7421133810e-01	8.8318132383e-01	-1.1308262390e+00	6.5841490091e-01	-1.1308262390e+00	1.2800000000e+04	5.8297327051e-03	9.8742187500e-01
26	-4.4991722330e-01	5.6895644492e-01	-1.1752933473e+00	7.2537612402e-01	-1.1752933473e+00	1.2800000000e+04	5.8226168506e-03	9.8500000000e-01
27	-4.4945662185e-01	6.1905943993e-01	-1.2082753720e+00	7.5881875015e-01	-1.2082753720e+00	1.2800000000e+04	5.8155699219e-03	9.8460937500e-01
28	-4.6614869027e-01	6.0342362639e-01	-1.1434005663e+00	6.7725187603e-01	-1.1434005663e+00	1.2800000000e+04	5.8298731104e-03	9.8773437500e-01
29	-4.4867885627e-01	6.6371999377e-01	-1.2089249899e+00	7.6024613359e-01	-1.2089249899e+00	1.2800000000e+04	5.8037394629e-03	9.8507812500e-01
30	-4.4291741659e-01	5.3073214410e-01	-1.1422381468e+00	6.9932073019e-01	-1.1422381468e+00	1.2800000000e+04	5.8237565967e-03	9.8523437500e-01
31	-4.5449746238e-01	6.2979144413e-01	-1.1714666786e+00	7.1696921626e-01	-1.1714666786e+00	1.2800000000e+04	5.8192064648e-03	9.8515625000e-01
32	-4.3343907379e-01	7.6529204960e-01	-1.2456427731e+00	8.1220369933e-01	-1.2456427731e+00	1.2800000000e+04	5.8202614502e-03	9.8492187500e-01
33	-4.5985057097e-01	7.3241162762e-01	-1.2082122116e+00	7.4836164066e-01	-1.2082122116e+00	1.2800000000e+04	5.8222529736e-03	9.8546875000e-01
34	-4.3731889909e-01	6.4833175066e-01	-1.2370337614e+00	7.9971486227e-01	-1.2370337614e+00	1.2800000000e+04	5.7991912793e-03	9.8500000000e-01
35	-4.5292956793e-01	6.1069023349e-01	-1.1986175002e+00	7.4568793225e-01	-1.1986175002e+00	1.2800000000e+04	5.8230514551e-03	9.8601562500e-01
36	-4.4224439493e-01	5.7917240857e-01	-1.2152456480e+00	7.7300125308e-01	-1.2152456480e+00	1.2800000000e+04	5.8231168018e-03	9.8492187500e-01
37	-4.4854262775e-01	1.827814051e+00	-1.2481932899e+00	7.9965066217e-01	-1.2481932899e+00	1.2800000000e+04	5.8055787256e-03	9.8398437500e-01
38	-4.4764457353e-01	8.1777756835e-01	-1.2155369633e+00	7.6789238977e-01	-1.2155369633e+00	1.2800000000e+04	5.8248659131e-03	9.8539062500e-01

Averaging MC quantities with qmca

```
Usage: qmca [options] [file(s)]
```

Options:

```
--version          show program's version number and exit
-v, --verbose      Print detailed information (default=False)
-q QUANTITIES, --quantities=QUANTITIES
                   Quantity or list of quantities to analyze. See names
                   and abbreviations below (default=all).
-u UNITS, --units=UNITS
                   Desired energy units. Can be Ha (Hartree), Ry
                   (Rydberg), eV (electron volts), kJ_mol (k.
                   joule/mole), K (Kelvin), J (Joules) (default=Ha).
-e EQUILIBRATION, --equilibration=EQUILIBRATION
                   Equilibration length in blocks (default=auto).
-a, --average      Average over files in each series (default=False).
-w WEIGHTS, --weights=WEIGHTS
                   List of weights for averaging (default=None).
-b, --reblock     (pending) Use reblocking to calculate statistics
                   (default=False).
-p, --plot        Plot quantities vs. series (default=False).
-t, --trace       Plot a trace of quantities (default=False).
-h, --histogram   (pending) Plot a histogram of quantities
                   (default=False).
-o, --overlay     Overlay plots (default=False).
--legend=LEGEND   Placement of legend. None for no legend, outside for
                   outside legend (default=upper right).
--noautocorr     Do not calculate autocorrelation. Warning: error bars
                   are no longer valid! (default=False).
--noac           Alias for --noautocorr (default=False).
--sac           Show autocorrelation of sample data (default=False).
--sv           Show variance of sample data (default=False).
-i, --image      (pending) Save image files (default=False).
-r, --report     (pending) Write a report (default=False).
-s, --show_options
                   Print user provided options (default=False).
-x, --examples   Print examples and exit (default=False).
--help          Print help information and exit (default=False).
```

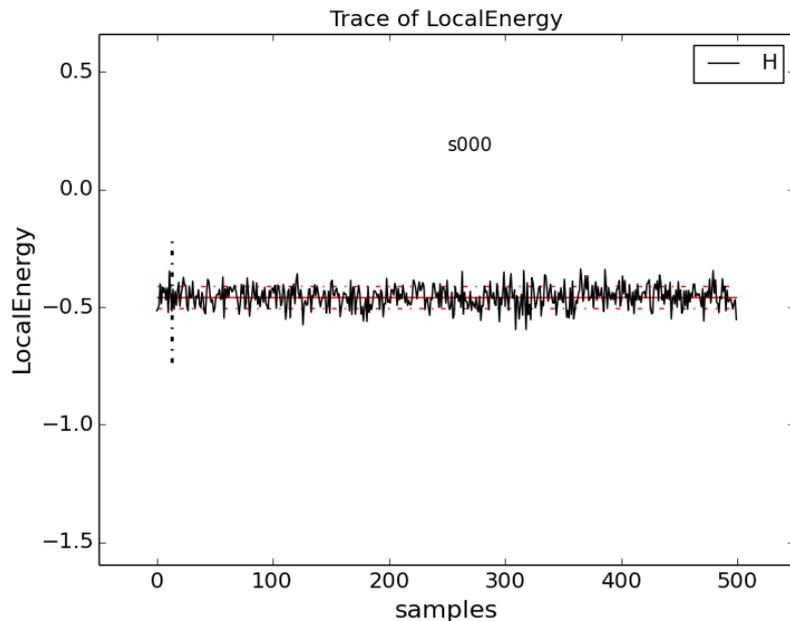
qmca

-q (quantities)

(prefix).scalar.dat

Evaluating simulation quality: trace, autocorrelation, and variance

Trace: `qmca -q e -t (prefix).scalar.dat`



Beware of large jumps or
regular oscillations!

Autocorrelation: `qmca -q e --sac (prefix).scalar.dat`

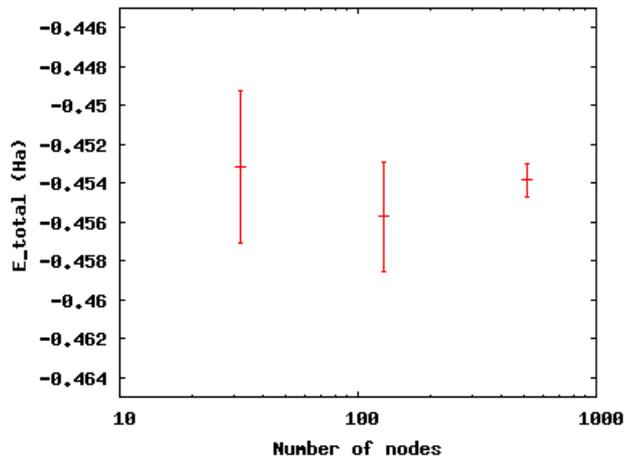
* Adjust time step and increase steps/block to lower

Variance: `qmca -q ev (prefix).scalar.dat`

* Aim for few percent variance : energy ratio

Reducing error bars: increased sampling, smaller variance

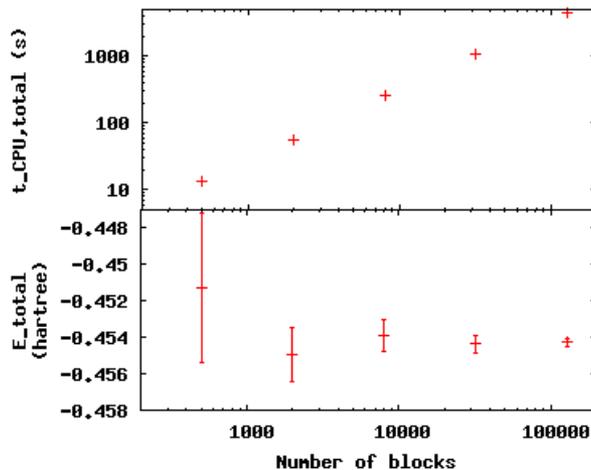
Number of nodes:



$$N_{\text{sample}} = N_{\text{node}} N_{\text{thread}} N_{\text{block}}$$

$$\sigma \sim \sqrt{\frac{1}{N_{\text{sample}}}}$$

Number of blocks:



*Quadrupling samples
halves error bars
(roughly)*

Variance:

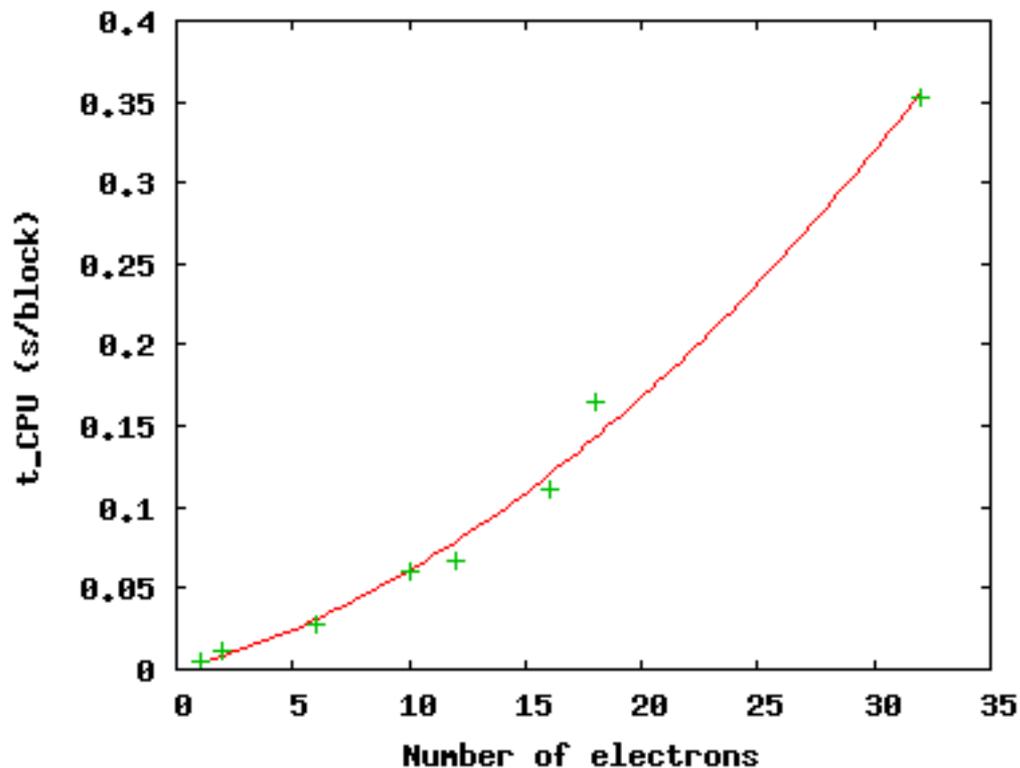
* Requires wave function optimization – Jastrow factor, basis set

Scaling the calculation up: measuring efficiency, increasing particle number

Efficiency: `qmca -q eff (prefix).scalar.dat` $\sim \frac{1}{t_{\text{CPU}} \sigma^2}$

Scaling with number of particles:

* $t_{\text{CPU}} \sim N^2$ for fixed basis set but σ increases too



Summary

1. Mean is sum of sample quantities over number of samples. Standard deviation is fluctuation measured by the square root of the sum of the second moment of the quantity's deviation from the quantity squared.
2. **QMCPACK** outputs MC data by block for into (prefix).scalar.dat by columns for quantities of interest
3. **qmca** averages quantities in (prefix).scalar.dat with error bars
4. Smooth trace, small autocorrelation, and vanishing variance indicate high quality MC simulation
5. Increasing number of samples and improving the wave function make for a better MC simulation
6. MC time scales with square of number of particles when basis is fixed.