

This document details the input parameters needed to reproduce all results in our forthcoming publication (<https://arxiv.org/abs/2005.00654>). More methodological details for each of these methods are provided in the paper. The statistics reported for each calculation can be obtained using the `calc_stats.py` Python script provided in this directory.

## FCI-FRI calculations

The following results were obtained using the systematic FCI-FRI method without the semi-stochastic extension, using the original heat-bath Power-Pitzer factorization.

### Ne (aug-cc-pVDZ)

These results are for the Neon atom in the aug-cc-pVDZ basis. The Hartree Fock data for this system, generated using `pyscf`, can be found in the directory `Input_Data/Neon_augccpvdz`. As an example, the results in the second line of this table were generated by running the following command:

```
mpirun -n 4 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/Neon_augccpvdz/ --target 100000 --distribution HB
--vec_nonz 100000 --mat_nonz 100000 --max_dets 50000 --initiator 0.5 --ini_vec Input_Data/Neon_augccpvdz/cisd_
```

Vec. nonz.	Matr. nonz.	Target <sup>1</sup>	$n_a^2$	Trial vec. <sup>3</sup>	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures <sup>4</sup>
100k	100k	100k	0.0	HF	CISD	40k	1M	$4.4561 \pm 16.5977$	0.0151	1
100k	100k	100k	0.5	HF	CISD	40k	1M	$2.3959 \pm 4.4132$	0.2139	1
100k	100k	100k	1.0	HF	CISD	40k	1M	$0.0223 \pm 0.0070$	85802	1, 2, 4
100k	100k	100k	1.5	HF	CISD	40k	1M	$0.0282 \pm 0.0064$	100190	1
100k	100k	100k	2.0	HF	CISD	40k	1M	$0.0321 \pm 0.0065$	97480	1
50k	50k	50k	1	HF	CISD	40k	1M	$0.0103 \pm 0.0095$	46368	2, 4
150k	150k	150k	1	HF	CISD	40k	1M	$0.0102 \pm 0.0055$	135837	2, 3, 4
200k	200k	200k	1	HF	CISD	40k	1M	$0.0052 \pm 0.0095$	206821	2, 3, 4
250k	250k	250k	1	HF	CISD	40k	1M	$0.0038 \pm 0.0040$	259087	2, 3
24k	52k	50k	3	HF	CISD	40k	1M	$0.0038 \pm 0.0101$	40798	5
42k	104k	100k	3	HF	CISD	40k	1M	$0.0196 \pm 0.0064$	100854	5
58k	157k	150k	3	HF	CISD	40k	1M	$0.0247 \pm 0.0051$	158244	5
73k	209k	200k	3	HF	CISD	40k	1M	$0.0251 \pm 0.0043$	229864	5

<sup>1</sup> Indicates the target one-norm of the iterates.

<sup>2</sup>  $n_a$  indicates the initiator threshold.

<sup>3</sup> Because no trial vector was specified in the input options, Hartree Fock was used.

<sup>4</sup> Indicates the figures in our paper in which the data appears.

## H<sub>2</sub>O (cc-pVDZ)

These results are for H<sub>2</sub>O in the cc-pVDZ basis. The Hartree Fock data for this system, generated using pyscf, can be found in the directory `Input_Data/H2O_ccpvdz`. As an example, the results in the second line of this table were generated by running the following command:

```
mpirun -n 16 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/H2O_ccpvdz/ --target 1000000 --distribution HB
--vec_nonz 1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 0.5 --ini_vec Input_Data/H2O_ccpvdz/cisd_
```

Vec. nonz.	Matr. nonz.	Target	$n_a$	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1M	1M	0.0	HF	CISD	40k	1M	$173.9337 \pm 49.0249$	0.00173	1
1M	1M	1M	0.5	HF	CISD	40k	1M	$2.4213 \pm 3.1589$	0.4176	1
1M	1M	1M	1.0	HF	CISD	40k	1M	$0.0171 \pm 0.0067$	92963	1,2
1M	1M	1M	1.5	HF	CISD	40k	1M	$0.0273 \pm 0.0064$	102492	1
1M	1M	1M	2.0	HF	CISD	40k	1M	$0.0298 \pm 0.0069$	86657	1
3M	3M	3M	1	HF	CISD	40k	524k	$0.0172 \pm 0.0048$	358168	2, 3, 4
5M	5M	5M	1	HF	CISD	40k	319k	$0.0124 \pm 0.0044$	732651	2, 3, 4
528k	1.1M	1M	3	HF	CISD	40k	1M	$0.0444 \pm 0.0059$	119583	5
3.4M	11.0M	10M	3	HF	CISD	40k	424k	$0.0174 \pm 0.0019$	2773910	5
9.9M	52.3M	50M	3	HF	CISD	40k	621k	$0.0155 \pm 0.0004$	42184887	5

## N<sub>2</sub> (cc-pVDZ)

These results are for N<sub>2</sub> in the cc-pVDZ basis. The Hartree Fock data for this system, generated using pyscf, can be found in the directory `Input_Data/N2_ccpvdz`. As an example, the results in the second line of this table were generated by running the following command:

```
mpirun -n 20 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/N2_ccpvdz/ --target 1000000 --distribution HB --vec_nonz
1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 0.5 --ini_vec Input_Data/N2_ccpvdz/cisd_
```

Vec. nonz.	Matr. nonz.	Target	$n_a$	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1M	1M	0.0	HF	CISD	40k	1M	$1593.6500 \pm 8414.6232$	5.88e-8	1
1M	1M	1M	0.5	HF	CISD	40k	1M	$284.2253 \pm 83.8154$	0.000593	1
1M	1M	1M	1.0	HF	CISD	40k	1M	$0.0275 \pm 0.0208$	9592	1,2
1M	1M	1M	1.5	HF	CISD	40k	1M	$0.0645 \pm 0.0161$	16103	1
1M	1M	1M	2.0	HF	CISD	40k	1M	$0.1103 \pm 0.0154$	17610	1
3M	3M	3M	1	HF	CISD	40k	1M	$0.0080 \pm 0.0101$	40632	2, 3, 4
5M	5M	5M	1	HF	CISD	40k	760k	$0.0087 \pm 0.0082$	83285	2, 3, 4
652k	1.2M	1M	3	HF	CISD	40k	1M	$0.1431 \pm 0.0127$	25676	5
4.8M	11.0M	10M	3	HF	CISD	40k	579k	$0.0167 \pm 0.0046$	348800	5
18.2M	53.4M	50M	3	HF	CISD	40k	272k	$0.0074 \pm 0.0022$	3684000	5

## Semi-stochastic FCI-FRI calculations

The following results were obtained using the systematic FCI-FRI method with the semi-stochastic extension, using the original heat-bath Power-Pitzer factorization.

### Ne (aug-cc-pVDZ)

These results are for the Neon atom in the aug-cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 4 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/Neon_augccpvdz/ --target 150000 --distribution HB
--vec_nonz 150000 --mat_nonz 150000 --max_dets 50000 --initiator 1.0 --ini_vec Input_Data/Neon_augccpvdz/cisd_ --det_space
Input_Data/Neon_augccpvdz/50_big_cisd_dets.txt
```

Nonz. <sup>1</sup>	$n_a^2$	Trial vec.	Ini vec.	Deterministic subspace <sup>3</sup>	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
150k	1	HF	CISD	50 largest CISD	40k	1M	$0.0113 \pm 0.0048$	179568	3
200k	1	HF	CISD	50 largest CISD	40k	1M	$0.0038 \pm 0.0039$	276718	3
250k	1	HF	CISD	50 largest CISD	40k	1M	$0.0041 \pm 0.0034$	351606	3
150k	1	HF	CISD	50 smallest CISD	40k	1M	$0.0146 \pm 0.0069$	86276	3
200k	1	HF	CISD	50 smallest CISD	40k	1M	$0.0042 \pm 0.0056$	134809	3
250k	1	HF	CISD	50 smallest CISD	40k	1M	$0.0015 \pm 0.0046$	351606	3

<sup>1</sup> Indicates the number of nonzero elements used in matrix and vector compressions. This number was also used as the target one-norm for all calculations in this section.

<sup>2</sup> Indicates the initiator threshold.

<sup>3</sup> The deterministic subspace was constructed from the Slater determinants corresponding to the largest- or smallest-magnitude elements in the CISD ground-state eigenvector.

## H<sub>2</sub>O (cc-pVDZ)

These results are for H<sub>2</sub>O in the cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 16 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/H2O_ccpvdz/ --target 1000000 --distribution HB
--vec_nonz 1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 1.0 --ini_vec Input_Data/H2O_ccpvdz/cisd_ --det_space
Input_Data/H2O_ccpvdz/150_big_cisd_dets.txt
```

Nonz.	$n_a$	Trial vec.	Ini vec.	Deterministic subspace	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1.0	HF	CISD	150 largest CISD	40k	1M	$0.0440 \pm 0.0203$	10107	3
3M	1.0	HF	CISD	150 largest CISD	40k	1M	$0.0135 \pm 0.0035$	344228	3
5M	1.0	HF	CISD	150 largest CISD	40k	661k	$0.0178 \pm 0.0029$	777820	3
1M	1.0	HF	CISD	150 smallest CISD	40k	1M	$0.0247 \pm 0.0102$	40239	3
3M	1.0	HF	CISD	150 smallest CISD	40k	590k	$0.0178 \pm 0.0054$	249487	3
5M	1.0	HF	CISD	150 smallest CISD	40k	642k	$0.0153 \pm 0.0042$	385820	3

## N<sub>2</sub> (cc-pVDZ)

These results are for N<sub>2</sub> in the cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 20 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/N2_ccpvdz/ --target 1000000 --distribution HB
--vec_nonz 1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 1.0 --ini_vec Input_Data/N2_ccpvdz/cisd_ --det_space
Input_Data/N2_ccpvdz/150_big_cisd_dets.txt
```

Nonz.	$n_a$	Trial vec.	Ini vec.	Deterministic subspace	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1.0	HF	CISD	150 largest CISD	40k	1M	$0.0408 \pm 0.0172$	14145	3
3M	1.0	HF	CISD	150 largest CISD	40k	1M	$0.0085 \pm 0.0072$	80267	3
5M	1.0	HF	CISD	150 largest CISD	40k	756k	$0.0159 \pm 0.0063$	141279	3
1M	1.0	HF	CISD	150 smallest CISD	40k	1M	$277.6687 \pm 52.1335$	0.0015	3
3M	1.0	HF	CISD	150 smallest CISD	40k	1M	$0.0306 \pm 0.0146$	19612	3
5M	1.0	HF	CISD	150 smallest CISD	40k	769k	$0.0141 \pm 0.0098$	57082	3

# FCI-FRI Calculations with the Alternative HB-PP Factorization

The following results were obtained using the systematic FCI-FRI method with the alternative, “un-normalized” HB-PP factorization, without the semi-stochastic extension.

## Ne (aug-cc-pVDZ)

These results are for the Neon atom in the aug-cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 4 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/Neon_augccpvdz/ --target 50000 --distribution HB_unnorm --vec_nonz 50000 --mat_nonz 50000 --max_dets 50000 --initiator 1.0 --ini_vec Input_Data/Neon_augccpvdz/cisd_
```

Nonz.	$n_a$	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
50k	1.0	HF	CISD	40k	1M	$0.0073 \pm 0.0079$	67108	4
100k	1.0	HF	CISD	40k	1M	$0.0114 \pm 0.0056$	132196	4
150k	1.0	HF	CISD	40k	1M	$0.0098 \pm 0.0042$	239832	4
200k	1.0	HF	CISD	40k	1M	$0.0048 \pm 0.0035$	333066	4

## H<sub>2</sub>O (cc-pVDZ)

These results are for H<sub>2</sub>O in the cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 16 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/H2O_ccpvdz/ --target 1000000 --distribution HB_unnorm --vec_nonz 1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 1.0 --ini_vec Input_Data/H2O_ccpvdz/cisd_
```

Nonz.	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1	HF	CISD	40k	1M	$0.0283 \pm 0.0058$	122875	4
3M	1	HF	CISD	40k	1M	$0.0184 \pm 0.0027$	583227	4
5M	1	HF	CISD	40k	1M	$0.0167 \pm 0.0023$	1148398	4

## N<sub>2</sub> (cc-pVDZ)

These results are for N<sub>2</sub> in the cc-pVDZ basis, using the same Hartree Fock data described above. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 20 ./build/FRIES_bin/frisys_mol --hf_path Input_Data/N2_ccpvdz/ --target 1000000 --distribution HB_unnorm --vec_nonz 1000000 --mat_nonz 1000000 --max_dets 100000 --initiator 1.0 --ini_vec Input_Data/N2_ccpvdz/cisd_
```

Nonz.	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	1	HF	CISD	40k	1M	$0.0697 \pm 0.0179$	13025	4
3M	1	HF	CISD	40k	1M	$0.0262 \pm 0.0080$	65667	4
5M	1	HF	CISD	40k	1M	$0.0154 \pm 0.0058$	146051	4

## FCIQMC (integer) Calculations

These results were obtained using our implementation of the FCIQMC method in which vector elements are constrained to be integers. The heat-bath Power-Pitzer matrix factorization was used in these calculations, along with the initiator approximation. The semi-stochastic extension was not used.

### Ne (aug-cc-pVDZ)

These results are for the Neon atom in the aug-cc-pVDZ basis, using the same Hartree Fock data described above. All calculations were initialized from a population of 5000 walkers distributed among determinants in proportion to the values of elements in the CISD unit vector. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 4 ./build/FRIES_bin/fciqmc_mol --hf_path Input_Data/Ne_augccpvdz/ --target 50000 --initiator 3 --max_dets 20000 --distribution HB --ini_vec Input_Data/Neon_augccpvdz/cisd_int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
50k	3	HF	CISD	40k	1M	$0.0099 \pm 0.0734$	773	5
100k	3	HF	CISD	40k	1M	$-0.0136 \pm 0.0546$	1400	5
150k	3	HF	CISD	40k	1M	$-0.0062 \pm 0.0457$	1998	5
200k	3	HF	CISD	40k	1M	$0.0178 \pm 0.0382$	2851	5

### H<sub>2</sub>O (cc-pVDZ)

These results are for H<sub>2</sub>O in the cc-pVDZ basis, using the same Hartree Fock data described above. These calculations were initialized from a population of 500,000 walkers distributed among determinants in proportion to the values of elements in the CISD unit vector. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 16 ./build/FRIES_bin/fciqmc_mol --hf_path Input_Data/H2O_ccpvdz/ --target 1000000 --initiator 3 --max_dets 100000 --distribution HB --ini_vec Input_Data/H2O_ccpvdz/cisd_int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	3	HF	CISD	40k	1M	$0.0491 \pm 0.0332$	3770	5
10M	3	HF	CISD	40k	574k	$0.0120 \pm 0.0147$	34817	5
50M	3	HF	CISD	40k	339k	$0.0147 \pm 0.0022$	2403490	5

## N<sub>2</sub> (cc-pVDZ)

These results are for N<sub>2</sub> in the cc-pVDZ basis, using the same Hartree Fock data described above. These calculations were initialized from a population of 500,000 walkers distributed among determinants in proportion to the values of elements in the CISD unit vector. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 8 ./build/FRIES_bin/fciqmc_mol --hf_path Input_Data/N2_ccpvdz/ --target 1000000 --initiator 3 --max_dets 200000 --distribution HB --ini_vec Input_Data/N2_ccpvdz/cisd.int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	3	HF	CISD	40k	1M	$0.1638 \pm 0.0781$	684	5
10M	3	HF	CISD	40k	543k	$0.0061 \pm 0.0398$	5010	5
50M	3	HF	CISD	40k	462k	$0.0089 \pm 0.0189$	24518	5

## FCIQMC (non-integer) Calculations

These results were obtained using our implementation of the FCIQMC method in which only some vector elements are integerized, and the rest are allowed to be non-integers. The heat-bath Power-Pitzer matrix factorization was used in these calculations, along with the initiator approximation. The semi-stochastic extension was not used.

## Ne (aug-cc-pVDZ)

These results are for the Neon atom in the aug-cc-pVDZ basis, using the same Hartree Fock data and initial vector as described in the previous FCIQMC section. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 4 ./build/FRIES_bin/fciqmc_fp_mol --hf_path Input_Data/Ne_augccpvdz/ --target 50000 --initiator 3 --max_dets 20000 --distribution HB --ini_vec Input_Data/Neon_augccpvdz/cisd.int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
50k	3	HF	CISD	40k	1M	$0.0307 \pm 0.0176$	13390	5
100k	3	HF	CISD	40k	1M	$0.0313 \pm 0.0126$	26240	5
150k	3	HF	CISD	40k	1M	$0.0336 \pm 0.0103$	39821	5
200k	3	HF	CISD	40k	1M	$0.0290 \pm 0.0090$	51644	5

## H<sub>2</sub>O (cc-pVDZ)

These results are for H<sub>2</sub>O in the cc-pVDZ basis, using the same Hartree Fock data and initial vector as described in the previous FCIQMC section. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 16 ./build/FRIES_bin/fciqmc_fp_mol --hf_path Input_Data/H2O_ccpvdz/ --target 1000000 --initiator 3 --max_dets 100000 --distribution HB --ini_vec Input_Data/H2O_ccpvdz/cisd_int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	3	HF	CISD	40k	1M	$0.0404 \pm 0.0097$	44495	5
10M	3	HF	CISD	40k	478k	$0.0165 \pm 0.0045$	442339	5
50M	3	HF	CISD	5k	334k	$0.0157 \pm 0.0021$	2818118	5

## N<sub>2</sub> (cc-pVDZ)

These results are for N<sub>2</sub> in the cc-pVDZ basis, using the same Hartree Fock data and initial vector as described in the previous FCIQMC section. As an example, the results in the first line of this table were generated by running the following command:

```
mpirun -n 8 ./build/FRIES_bin/fciqmc_fp_mol --hf_path Input_Data/N2_ccpvdz/ --target 1000000 --initiator 3 --max_dets 200000 --distribution HB --ini_vec Input_Data/N2_ccpvdz/cisd_int_
```

Target walkers	Ini. thresh. ( $n_a$ )	Trial vec.	Ini vec.	Burn-in	Iterations	Mean $\pm 2\sigma$ ( $mE_h$ )	Efficiency ( $E_h^{-2}$ )	Figures
1M	3	HF	CISD	40k	1M	$0.1316 \pm 0.0190$	10574	5
10M	3	HF	CISD	40k	438k	$0.0095 \pm 0.0099$	103387	5
50M	3	HF	CISD	40k	307k	$0.0097 \pm 0.0060$	422516	5