

# Rank-Restricted Full Configuration Interaction

Nick F. Settje and James W. Snyder Jr

## I. Introduction

Configuration interaction is a widely used method in electronic structure theory for optimizing a wavefunction by minimizing the energy of the time-independent electronic Schrödinger equation, which is the standard way to determine the energy of a quantum system. The wavefunction, which determines all properties of a quantum system, can be expressed as,

$$|\Psi\rangle = \sum_I c_I |\Phi_I\rangle,$$

which is a linear combination of n-electron configurations, or determinants, with coefficients equal to  $c_I$ . The components  $|\Phi_I\rangle$  form an orthonormal basis of the wavefunction. The determinants are the set of features that comprise the wavefunction, so we can optimize them by minimizing the electronic energy. To determine the energy of the wavefunction  $c$ , the Schrödinger equation in matrix form reads

$$Hc = Ec,$$

where  $H$  is the Hamiltonian matrix with elements  $H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle$ ,  $E$  is a vector containing the energy of each configuration, and  $c$  is the vector containing the CI coefficients. A detailed account of the origin of the determinants that comprise  $c$  as well as that of the Hamiltonian operator,  $\hat{H}$ , are beyond the scope of this paper. However, it is important to note that each matrix element of  $H$  is an integral that can be readily computed.

If the above expansion includes every possible configuration of electrons allowed given a basis set of finite size, then this is known as full configuration interaction (FCI). FCI is the best possible approximation to the Schrödinger equation. As the number of determinants approaches infinity, FCI will approach the exact solution to the Schrödinger equation. Because one must compute every possible configuration of electrons, the above eigenvalue problem scales as  $O(n!)$ , since there are  $n!$  configurations of  $n$  electrons. This means FCI is intractable for anything besides trivially small systems.

For additional information on the fundamentals of electronic structure theory, refer to [1] and [2]. For a thorough treatment of FCI, consult [3].

## II. Solving Full CI by Gradient Descent

Our implementation of FCI is based upon the graphical unitary group approximation for string based FCI [3]. The full details of this implementation are largely beyond the scope of this paper. In simple terms, the standard method is to compute various integrals, build the matrix elements of our Hamiltonian  $H$ , evaluate the product  $Hc$ , and iterate over this procedure until convergence of the  $c$  vector for each eigenstate of  $H$ . Because  $H$  is a matrix of dimension  $n!^2 \times n!^2$  we build  $Hc$  implicitly without ever representing  $H$ . Even without these extra storage requirements, FCI remains exceedingly expensive.

Luckily, optimization of the FCI wavefunction benefits greatly from machine learning algorithms. The wavefunction coefficients can be optimized using gradient descent. Since the objective of FCI is to minimize the electronic energy of the wavefunction, the gradient descent rule for FCI is

$$|c_k\rangle := |c_k\rangle - \nabla_{c_k} E,$$

where

$$E = \frac{\sum_{ij} c_i c_j \langle \Phi_i | \hat{H} | \Phi_j \rangle}{\sum_{ij} c_i c_j \langle \Phi_i | \Phi_j \rangle}.$$

If one takes the gradient with respect to  $c_k$ ,

$$\begin{aligned} \nabla_{c_k} E &= \frac{\sum_j c_j \langle \Phi_k | \hat{H} | \Phi_j \rangle}{\sum_{ij} c_i c_j \langle \Phi_i | \Phi_j \rangle} - \frac{(\sum_{ij} c_i c_j \langle \Phi_i | \hat{H} | \Phi_j \rangle) \sum_j c_j \langle \Phi_k | \hat{H} | \Phi_j \rangle}{(\sum_{ij} c_i c_j \langle \Phi_i | \Phi_j \rangle)^2} \\ \nabla_{c_k} E &= \frac{|\sigma_k\rangle}{\sum_i c_i^2} - \frac{(\sum_{ij} \langle c_i | \sigma_j \rangle) |c_k\rangle}{(\sum_i c_i^2)^2} \end{aligned}$$

where  $\sigma = \hat{H}c$ . This gives us the update rule,

$$|c_k\rangle := |c_k\rangle - \frac{|\sigma_k\rangle}{\sum_i c_i^2} + \frac{(\sum_{ij} \langle c_i | \sigma_j \rangle) |c_k\rangle}{(\sum_i c_i^2)^2}.$$

Since  $\sum_i c_i^2 = n$  for a normalized wavefunction, this update rule can be rewritten as

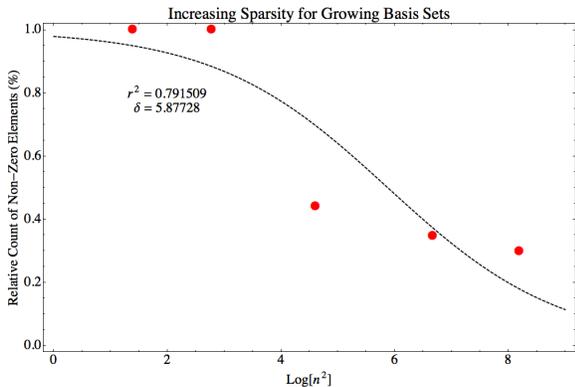
$$|c_k\rangle := \left(1 + \frac{\sum_{ij} \langle c_i | \sigma_j \rangle}{n^2}\right) |c_k\rangle - \frac{|\sigma_k\rangle}{n},$$

where  $n$  is the number of eigenstates being solved for. The update rule only depends on  $c$  and  $\sigma$ , so this method does not require extra memory beyond what we would already use to represent these vectors. This is especially beneficial because of the difficult scaling of FCI, where extra intermediates simply would not fit in core for larger molecules.

Admittedly, better methods for converging the  $c$  vectors exist than simple gradient descent. However, all of these methods introduce extra structure into the  $c$  vectors after each iteration, as they involve diagonalization of sub-Hamiltonians in subspaces of the set of determinants. This makes no difference in regular FCI, but it can cause problems if we try to factor our  $c$  vectors to reduce the scaling of the problem.

### III. Rank-Restricted Full CI

We justify the use of gradient descent by observing that the converged  $c$  vectors for FCI are exceedingly sparse. This motivates exploiting the string based method for writing  $c$  vectors as matrices



**Fig. 1:** Logistic fit of relative sparsity for increasing number of determinants in FCI for  $H_2$ . The sudden drop in non-zero elements motivates factorizing  $c$ .

indexed by  $\alpha$  (up spin) and  $\beta$  (down spin) electrons. As shown in Figure 1, the relative sparsity (percentage of non-zero matrix elements of  $c$  compared to the set of determinants with the densest  $c$  matrix) of such matrices decreases as the size of  $c$  increases. In this plot, we see a sharp drop in non-zero elements near the phase transition  $\delta$ .

The sparsity of such a matrix allows for a Cholesky-like decomposition:

$$c_{\alpha,\beta} \approx \sum_p D_{\alpha,p} D_{\beta,p}$$

where  $p \ll \alpha, \beta$ . For molecules with the same number

of  $\alpha$  and  $\beta$  electrons,  $c$  is a square matrix. Thus, the Schrodinger equation becomes

$$Hc = HDD^T = \sigma$$

Effectively, this factorization asserts that the structure of  $c$  must depend largely on only a small number of rows or columns, which we write as the rectangular  $D$  matrices. In the language of machine learning, rank restriction is feature reduction. We assert that the bias introduced from considering a subset of the features is negligible with respect to the accuracy of the factored model. Fortunately, the structure of the  $c$  matrices allows for decomposition without sacrificing much accuracy. This now motivates us to derive some iterative scheme for converging  $D$  instead of  $c$ . Since we require that  $D$  obey certain structure, we want an algorithm that offers control over the imposed structure after each iteration. Gradient descent is just such an algorithm.

In terms of  $D$ , we write

$$\mathcal{E} = \frac{\sum_{ij} D_i D_j \langle \Phi_i | \hat{H}_{red} | \Phi_j \rangle}{\sum_{ij} D_i D_j \langle \Phi_i | \Phi_j \rangle},$$

where

$$H_{red}D = \mathcal{E}D.$$

Given this, the gradient descent rule for low rank FCI can be written as,

$$|D_k\rangle := |D_k\rangle - \nabla_{D_k} \mathcal{E}.$$

Because this takes the exact same form as the full rank gradient descent rule, the solution is evident upon inspection.

$$|D_k\rangle := \left( 1 + \frac{\sum_{ij} \langle D_i | \sigma_{red,j} \rangle}{n^2} \right) |D_k\rangle - \frac{|\sigma_{red,k}\rangle}{n}.$$

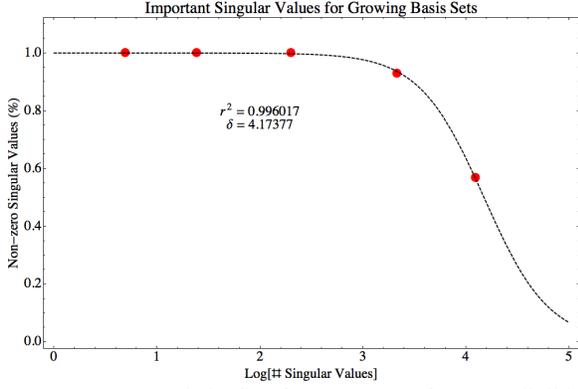
These two iterative update algorithms constitute a set of machine learning algorithms. Although these algorithms specifically optimize the wavefunction to minimize the energy of the system, the molecular wavefunction actually contains all information related to a molecular system. Subsequent to performing this optimization, it is possible to predict a variety of properties, from the dipole moment to the ionization potential, using the set of optimized determinants, which are the features used to optimize the energy. From the decomposition of  $c$ , we can write the full energy in terms of  $\mathcal{E}$

$$Hc = (H_{red}D)^2 = Ec = (\mathcal{E}D)^2$$

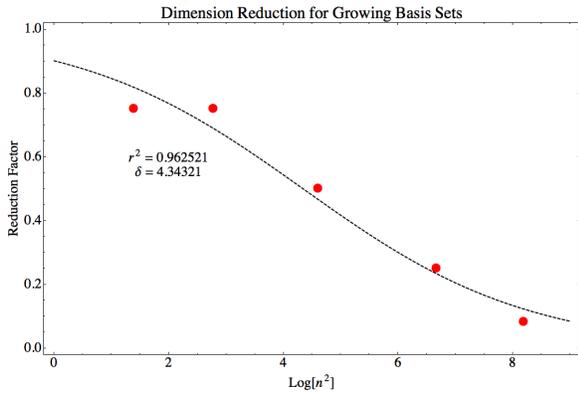
So we only have to square the reduced energy to recover the energy of the full system.

#### IV. Results

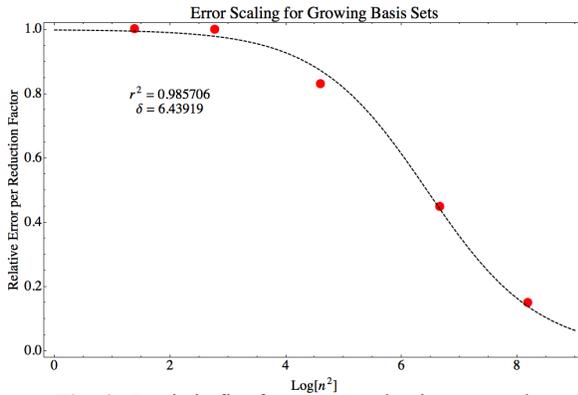
In order to perform rank-restricted FCI, it is necessary to select a value for  $p$  for use in building  $D$ . The goal is to choose a value of  $p$  large enough to capture almost all of the non-zero elements of  $c$  but small enough to give a significant speed up in calculation. As a rough measure of the information content of  $c$ , we perform singular value decomposition on the  $c$  matrix and retain only as many rows or columns of  $c$  as the number of singular values above some threshold. As shown in Figure 2, the percentage of non-negligible singular values decreases rapidly as the number of determinants increases. To determine the extent to which this decreases the number of matrix elements we need to retain, we calculate the reduction factor as the percentage of important singular values. This is the factor by which the number of determinants would be multiplied to find the total number we need to retain. As Figure 3 shows, the



**Fig. 2:** Logistic fit of percentage of non-negligible singular values as number of determinants increase in FCI for  $H_2$ . The sudden drop in important values further motivates factorizing  $c$ .



**Fig. 3:** Logistic fit of percentage of matrix elements that probably must be retained as number of determinants increase in FCI for  $H_2$ . The sudden drop in number of retained elements further motivates factorizing  $c$ .



**Fig. 4:** Logistic fit of error per reduction as number of determinants increase in FCI for  $H_2$ . The phase transition gives a lower bound on the value of  $p^*$ . The error decreases suddenly and approaches zero for large enough sets of determinants.

that modern calculations use more than  $10^6$  determinants.

Since on average we use roughly the square root of the number of determinants for a given factorization, the scaling of the FCI algorithm becomes  $O(\sqrt{n!})$  for approximations to within the convergence tolerance of the full rank algorithm. This scaling may still seem rather abysmal, but in the

reduction factor reduces quickly as the number of determinants increases. This means that we must retain a smaller fraction of the determinants as their number increases.

In order to determine the extent to which these factorizations affect the energy, we compare the energy of the rank-restricted calculation to that for the full rank computation. Since we are interested in the effect of restricting the size of  $D$  on the total error, a natural measure of the efficacy of our model is the absolute error between the full and restricted models divided by the reduction factor

$$J(p) = \frac{|E_{full} - \epsilon_p|^2}{\gamma}$$

where  $\gamma$  is the reduction factor. Intuitively, this measures how well our model does as a function of the rank restriction. This is apparent when we notice

$$\lim_{p \rightarrow \infty} J(p) = \lim_{\gamma \rightarrow 1} J(p) = 0$$

because  $D$  approaches  $c$  as the number of rows retained approaches the entirety of  $c$ , which means that the reduction factor approaches 1, since there is no reduction in the number of matrix elements retained. This means that we are minimizing the cost function  $J(p)$  as well as the reduction factor  $\gamma$

$$p^* = \operatorname{argmin}_p J(p) \text{ s.t. } p = n^2 \gamma$$

where  $p^*$  is the optimal  $p$  for reducing the number of matrix elements retained and approximating the total energy with the best accuracy. The constraint on  $p$  follows from the fact that  $c$  has  $n^4$  matrix elements, so  $D$  has  $n^2 p$  elements, so the total reduction factor is  $\gamma = \frac{p}{n^2}$ .

As shown in Figure 4, the absolute energy difference per reduction factor  $J(p)$  decreases as the number of determinants increases. This means that the rank-restricted approximation has smaller error per reduction for larger matrices that capture more of the electronic energy. The phase transition  $\delta$  is a lower bound on  $p^*$ . Because of this, we expect to factor more aggressively starting at calculations using more than  $\sim 2000$  determinants. This is rather small considering

context of chemistry it is the difference between calculating on water and calculating on the active site of a biologically important protein.

## V. Future Work

The results of this study focused on FCI calculations of the small test molecule  $H_2$ . The most important result to check now is whether the phase transition of factorization is independent of the number of electrons. This is most easily accomplished by repeating this analysis for a larger set of molecules with more electrons, such as the traditional test set of HF, BH,  $H_2O$ , and  $CH_4$ . If the factorization approaches the same optimal value of  $p$  for all systems, this method would be extremely useful for calculating FCI-like energies for molecules that are too large for regular FCI.

Since gradient descent requires a large number of iterations to converge, it is somewhat slow for running rank-restricted calculations. Future implementations of rank restriction might consider choosing a Newton-Raphson update of the form

$$c_k := c_k - \frac{\nabla_{c_k E}}{\nabla_{c_m c_k}^2} = c_k - \frac{n^2 \nabla_{c_k E}}{2E}$$

since the Hessian adds extra information about the topology of the energy surface without introducing unwanted structure or reducing sparsity in the updated vectors. In some sense, the factor appearing before the gradient is the optimal learning rate, so the curvature of the energy surface determines the speed with which the algorithm converges.

Moreover, factoring  $c$  into two  $D$  matrices is the simplest rank restriction but it may not be the most accurate or the fastest to converge. Future studies might consider other decompositions such as LDL, in which the  $c$  matrices become

$$c_{\alpha,\beta} \approx \sum_{p,q} L_{\alpha,p} D_{p,q}^{\frac{1}{2}} D_{p,q}^{\frac{1}{2}} L_{q,\beta} = \sum_{p,q} \left( L_{\alpha,p} D_{p,q}^{\frac{1}{2}} \right)^2$$

where  $L$  is lower triangular and  $D$  is diagonal. Though we must now optimize two matrices at the same time, it may be that we must store and calculate fewer matrix elements overall to approximate regular FCI to high order.

## VI. References

- [1] Szabo, A. and Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola: Dover Publishing Inc., 1996.
- [2] Helgaker, T., Jørgensen, P., & Olsen, J. *Molecular electronic-structure theory*. Chichester: Wiley, 2000.
- [3] Sherrill, C. D. and Schaefer, H. F. *The Configuration Interaction Method: Advances in Highly Correlated Approaches*. Advances in Quantum Chemistry, Academic Press, 1999, Volume 34, Pages 143-269.
- [4] Turney, J. M., Simmonett, A. C., Parrish, R. M., Hohenstein, E. G., Evangelista, F., Fermann, J. T., Mintz, B. J., Burns, L. A., Wilke, J. J., Abrams, M. L., Russ, N. J., Leininger, M. L., Janssen, C. L., Seidl, E. T., Allen, W. D., Schaefer, H. F., King, R. A., Valeev, E. F., Sherrill, C. D., and Crawford, T. D. *Psi4: An Open-Source ab initio Electronic Structure Program*. *WIREs Computational Molecular Science*. **2**, 556 (2012).