Computational Scaling of the Configuration Interaction Method with System Size

C. David Sherrill

Department of Chemistry

University of California, Berkeley

September 1996

At times there appears to be a confusion in the literature regarding the computational scaling of the configuration interaction method. For example, Gershgorn and Shavitt claim [1] that CI scales as $\mathcal{O}(n^{m+2})$, where n is the number of orbitals, and m is the excitation level (e.g., 2 for CISD). However, it is commonly held that CISD actually scales as $\mathcal{O}(n^6)$; for CISDTQ, various scalings are given, but the most common one is $\mathcal{O}(n^{10})$. This would seem to imply a general scaling rule of $\mathcal{O}(n^{2m+2})$, in contradiction to the former rule. These notes explain the scaling of the CI method and demonstrate that, depending on circumstances, either limit may be appropriate. For the present purposes, it will suffice to focus simply on Slater determinants rather than other possible N-electron basis functions, such as configuration state functions (CSF's).

For present purposes, it is sufficient to work with spin orbitals. Typically, the dimension of the CI space is dominated by determinants with the maximum excitation level, m. Thus,

$$N_{det} \sim \left(\begin{array}{c} N \\ m \end{array} \right) \left(\begin{array}{c} n_v \\ m \end{array} \right) \sim \frac{1}{(m!)^2} N^m n_v^m, \tag{1}$$

with n_v orbitals unoccupied in the reference (i.e., virtual orbitals). Most CI procedures solve only for the lowest or lowest few eigenvectors, via an iterative procedure. In such situations, the scaling is much less than the $\mathcal{O}(N_{det}^3)$ typical of standard matrix diagonalization methods. The most expensive step in iterative procedures such as the Davidson method [2] is the construction of the so-called σ vectors,

$$\sigma_i = \mathbf{H}\mathbf{b}_i \tag{2}$$

where $\mathbf{b}_i$ belongs to a set of trial vectors which is expanded each iteration until convergence is reached. If the Hamiltonian matrix $\mathbf{H}$ were formed directly, this procedure would require $\mathcal{O}(N_{det}^2)$ operations. This is never actually done because the storage requirements would be too great, and such an approach ignores the fact that the Hamiltonian contains only two-body terms, so that the majority of the matrix elements are zero.

Each element of a trial vector $\mathbf{b}_i$ need only be multiplied by the nonzero elements of $\mathbf{H}$. The Hamiltonian will connect a maximally excited determinant with other maximally excited

determinants and with other determinants having excitation level $m \geq m' \geq m-2$. The number of interacting determinants is roughly

$$\begin{pmatrix} m \\ 2 \end{pmatrix} \begin{pmatrix} n_v + m \\ 2 \end{pmatrix} + \begin{pmatrix} m \\ 2 \end{pmatrix} \begin{pmatrix} N - m \\ 2 \end{pmatrix} + m^2 n_v (N - m), \tag{3}$$

which we further approximate as

$$\frac{1}{4} m^2 (n_v + m)^2 + \frac{1}{4} m^2 (N - m)^2 + m^2 n_v (N - m). \tag{4}$$

Each element in $\mathbf{b}_i$ must be multiplied by the relevant nonzero matrix elements, leading to an overall operation count on the order of

$$\mathcal{O}(N^m n_v^m \left\{ m^2 n_v^2 + m^2 N^2 + m^2 N n_v \right\}). \tag{5}$$

Except for full CI, we typically expect $N, n_v >> m$. Furthermore, we almost always have $n_v > N$, so that the leading term becomes $\mathcal{O}(N^m n_v^{2+m})$. Thus the final results are

$$\mathcal{O}(N^m n_v^{2+m}) = \begin{cases} n^{2m+2} & \text{if } N \sim n_v >> m \\ n^{m+2} & \text{if } n_v >> N, m. \end{cases} \tag{6}$$

These results show that CISD and CISDTQ will scale as $\mathcal{O}(n^6)$ and $\mathcal{O}(n^{10})$, respectively, for typical cases, and as $\mathcal{O}(n^4)$ and $\mathcal{O}(n^6)$ for very large numbers of orbitals or a small number of active electrons.

For very high levels of excitation, including full CI, the number of interacting matrix elements for a given determinant becomes approximately $N^2 n^2$, so that the computational cost becomes roughly

$$\mathcal{O}^{FCI}(N_{det} N^2 n^2) \tag{7}$$

where we have replaced the term $N^m n_v^m$ with the actual number of determinants, N_{det}. For comparison, the determinant full CI algorithm of Knowles and Handy [3] scales as $\mathcal{O}(N_{det} n^4)$, while the algorithm of Olsen et al. [4] and similar approaches scale as $\mathcal{O}(N_{det} N^2 (n - N/2)^2)$ for a closed-shell system.

References

[1] Z. Gershgorn and I. Shavitt, Int. J. Quantum Chem. **2**, 751 (1968).

[2] E. R. Davidson, J. Comp. Phys. **17**, 87 (1975).

[3] P. J. Knowles and N. C. Handy, Chem. Phys. Lett. **111**, 315 (1984).

[4] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen, J. Chem. Phys. **89**, 2185 (1988).