



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS

DESARROLLO DE UN MÉTODO PARA CÁLCULO DE ESTRUCTURAS ELECTRÓNICAS DE MOLÉCULAS PEQUEÑAS (HASTA 50 ELECTRONES) Y SUS APLICACIONES.

TESIS
QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. en C. César Xavier Almora Díaz

TUTOR PRINCIPAL

Dr. Carlos F. Bunge Molina
Instituto de Física

MIEMBROS DEL COMITÉ TUTOR

Dr. Ilya Kaplan Savitsky
Instituto de Investigaciones en Materiales

Dr. Fernando Colmenares Landín
Facultad de Química

MÉXICO, D. F. FEBRERO 2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Jurado asignado por el Comité Académico para la evaluación del trabajo de tesis de Doctorado:

Presidente:

Dr. Eugenio Ley-Koo Instituto de Física, UNAM

Vocal:

Dr. Emilio Orgaz Baqué. Facultad de Química, UNAM

Vocal:

Dr. Alejandro Ramírez Solís. Facultad de Ciencias, UAEMorelos.

Vocal:

Dr. Raúl Álvarez Idaboy Facultad de Química, UNAM

Secretario:

Dr. Fernando Colmenares Landín. Facultad de Química, UNAM

*A mis dos amores:
Mathias y Élise*

Agradecimientos

A la Universidad Nacional Autónoma de México.

Al CONACyT por la beca de estudio de posgrado número 216080.

Al Instituto de Física de la UNAM.

Al Posgrado en Ciencias Químicas. Especialmente a la Dra. Adela Rodríguez Romero y a la Q.F.B. Josefina Tenopala García.

A Carlos F. Bunge por tu apoyo, confianza, amistad y respecto a lo largo de tantos años, por formarme como científico y hacerme una mejor persona.

A los miembros de Comité tutorial el Dr. Ilya Kaplan G. Kaplan y el Dr. Fernando Colmenares Landín.

A los miembros de mi Jurado de Tesis. Especialmente al Dr. Eugenio Ley-Koo, muchas gracias por tenerme tanta paciencia.

Al Dr. Jeff Hammond, por explicarme varios de los secretos de NWChem y por mostrarme parte de tu trabajo de doctorado gracias al cual fue posible terminar el mío, por toda tu amabilidad y atención.

Al personal del servicio de cómputo académico del Instituto de Física y a la Dirección General de Tecnologías de la Información y Cómputo, especialmente a José Luis Gordillo.

Gracias a mi Padre y a mi Madre.

A toda mi familia.

Éste trabajo fue realizado en el Instituto de Física de la Universidad Nacional Autónoma de México. Utilizando los recursos de cómputo otorgados por el servicio de cómputo académico del Instituto de Física y de la Dirección General de Tecnologías de la Información.

Los resultados de éste trabajo fueron presentados en los siguientes congresos.

- 1.- QuimiUNAM 2013, Ciudad Universitaria, México D.F., México (póster).
- 2.- XII Reunión de la Sociedad Mexicana de Fisicoquímica Teórica, Juriquilla, Querétaro, México, Noviembre 2013.
- 3.- 26th Canadian Symposium on Theoretical and Computational Chemistry, Montréal, Quebec. Canadá, Julio, 2014 (póster).
- 4.- Concepts of Mathematical Physics in Chemistry Workshop en honor del Dr. Frank E. Harris, Playa del Carmen, México, Diciembre, 2014.

Los resultados del Capítulo 5 de ésta tesis fueron publicados en el artículo:

”Highly correlated configuration interaction calculations on water with large orbital bases”, César X. Almora-Díaz, *The Journal of Chemical Physics* 140, 184302 (2014); doi: 10.1063/1.4874319.

Resumen

Esta tesis contiene los resultados de la primera implementación del método de interacción de configuraciones con error de truncamiento y análisis de sensibilidad (ICSET) [1] y de interacción de configuraciones seleccionado y por partes (ICSPP) [2] aplicados a sistemas moleculares. Estos métodos tienen como objetivo la obtención de espacios modelo que permitan el cálculo de cotas superiores a la energía que aproximen los valores propios obtenidos con los métodos de interacción de configuraciones (IC) convencionales.

En el Capítulo 1 se lleva a cabo la presentación de los métodos de estructura electrónica implementados en éste trabajo, su desarrollo histórico y algunos de los principales resultados obtenidos.

En el Capítulo 2 se presentan antecedentes generales de los fundamentos de los métodos de estructura electrónica de primeros principios, con un enfoque específico en los fundamentos del método de Interacción de Configuraciones y se hace una breve comparación con otros métodos de primeros principios. En los Capítulos 3 y 4 se presentan los métodos de estructura electrónica utilizados y desarrollados en éste trabajo.

El Capítulo 5 muestra los primeros resultados de nuestra implementación de los métodos presentados en los anteriores capítulos, comenzando con la validación de dichas implementaciones y posteriormente se presentan los mejores resultados variacionales reportados en la literatura para la energía total no relativista de la molécula de H₂O en la geometría de equilibrio de estado fundamental.

El Capítulo 6 contiene un pequeño estudio de la disociación simétrica de la molécula de H₂O con el cual fue posible evaluar el rendimiento de nuestros métodos con respecto a los mejores métodos multirreferenciales actuales. Y se explica una variante de dichos métodos que es capaz de producir los mejores resultados variacionales, para éste sistema.

El Capítulo 7 presenta una serie de resultados para un conjunto de oligómeros de átomos idénticos separados a distancia infinita que muestran el desempeño de los métodos utilizados en éste trabajo con respecto al denominado problema de consistencia en tamaño.

Finalmente, en el Capítulo 8 se presentan las conclusiones generales de éste trabajo.

Abstract

This thesis contains the results of the first implementation of configuration interaction with truncation error and sensitivity analysis, and of selected configuration interaction by parts methods applied to molecular systems. These methods aim at obtaining model spaces which allow the calculation of upper bounds that approximate the values obtained with conventional configuration interaction configurations methods.

In Chapter 1, the electronic structure methods implemented in this work are introduced, altogether with its historical development and some of the principal results obtained in this work.

In Chapter 2, general background of the foundations of first principal methods of electronic structure is presented, with a specific focus on the fundamentals of Configuration Interaction method and a brief comparison with other first principles methods.

In Chapters 3 and 4 electronic structure methods used and developed in this work are presented.

Chapter 5 shows the first results of our implementation of the methods presented in Chapters 3 and 4, beginning with the validation of these implementations and subsequently the best variational results of the nonrelativistic total energy for the ground state of the water molecule.

Chapter 6 contains a small study of the symmetric dissociation of water molecule with which it was possible to compare the performance of our methods with respect to the best available multireferential results. A variant of these methods that is capable of producing the best variational results for this system is explained.

Chapter 7 presents results for a set of oligomers of identical atoms separated at "infinite" distance that show the performance of the methods used throughout this work with respect to the problem of consistency in size of truncated configuration interaction methods.

Finally, in Chapter 8 the general conclusions of this work are presented.

Índice general

Agradecimientos	III
Resumen	VI
Índice general	VIII
Índice de figuras	X
Índice de tablas	XI
Abreviaciones	XIII
1. Introducción	1
2. Antecedentes generales de Interacción de Configuraciones	6
2.1. Ecuación de Schrödinger independiente del tiempo para moléculas.	6
2.2. Interacción de Configuraciones	7
2.3. Funciones de simetría de n electrones.	9
2.4. Orbitales naturales.	10
2.5. Error de truncamiento de la base.	11
2.6. Interacción de Configuraciones frente a otros métodos de primeros principios	11
3. Interacción de Configuraciones seleccionado.	13
3.1. Antecedentes.	13
3.1.1. Truncamiento por nivel de excitación.	13
3.1.2. Selección de determinantes bajo criterios de número de ocupación.	14
3.1.3. Selección bajo criterios de espacios interactuantes.	15
3.1.4. Selección bajo criterios de energía y de coeficientes de IC.	15
3.2. Interacción de Configuraciones con errores de truncamiento.	17
3.2.1. La fórmula de Brown.	18
3.2.2. Desarrollo de cúmulos y predicción de coeficientes de expansión configuracionales.	20
3.2.3. Criterio de selección para configuraciones conectadas.	24
4. Interacción de Configuraciones seleccionado y por partes.	26
4.1. Criterios de selección de los subespacios de S.	28

4.1.1. Criterios de Energía y pseudoenergía.	28
5. Estado fundamental del H₂O.	30
5.1. Antecedentes	30
5.2. Resultados y discusión.	31
5.3. Conclusiones.	36
6. Disociación simétrica del H₂O.	38
6.1. Antecedentes	38
6.2. Resultados y discusión	39
6.3. Comparación con métodos multirreferenciales.	49
6.4. Conclusiones.	53
7. Consistencia en tamaño y algunos ejemplos.	54
7.1. Antecedentes.	54
7.2. Resultados y discusión.	55
7.3. Conclusiones	60
8. Conclusiones	61
Bibliografía	63

Índice de figuras

6.1. Espacios de IC	48
6.2. Comparación 1	50
6.3. Comparación 2	50
6.4. Multirreferencial 1	51
6.5. Multirreferencial 2	51

Índice de tablas

5.1. Energías totales (en Hartree) para el estado fundamental de H ₂ O en una bases RADZ	32
5.2. Energías SCF para el H ₂ O en estado fundamental.	33
5.3. Energías de correlación para el H ₂ O en estado fundamental (Hartree) y los valores extrapolados, para la familia de bases orbitales cc-pCVnZ	35
5.4. Energías de correlación para el H ₂ O en estado fundamental (Hartree) y los valores extrapolados, para la familia de bases orbitales aug-cc-pCVnZ	35
5.5. Contribución a la energía de correlación para el H ₂ O en estado fundamental (Hartree).	36
6.1. Energías totales en Hartree para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados por Olsen <i>et al.</i> [3]	39
6.2. Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.1	40
6.3. Energías totales de IC directo (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos usando PSI4.	41
6.4. Energías totales de ICS (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSD	41
6.5. Energías totales de ICS (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDT	42
6.6. Energías totales de ICS (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDTQ	42
6.7. Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.3.	42
6.8. Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.4.	42
6.9. Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.5.	43

6.10. Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H ₂ O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.6.	43
6.11. Tamaños de los subespacios S ₀ , S ₀₁ y los espacios totales en los cuales se calcularon los ICS hasta séxtuplos y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF reportados por Olsen <i>et al.</i>	45
6.12. Tamaños de los subespacios S ₀ , S ₀₁ y los espacios totales en los cuales se calcularon los ICS hasta séxtuplos y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDT.	45
6.13. Tamaños de los subespacios S ₀ , S ₀₁ y los espacios totales en los cuales se calcularon los ICS hasta séxtuplos y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDTQ.	46
6.14. Diferencias de energía con respecto al ICC (en Hartree) para el estado fundamental de H ₂ O en una base cc-pVDZ a distintas distancia H-O, con distintos métodos multirreferenciales y algunos resultados de las Tablas 6.2, 6.6, 6.11 y 6.12.	52
7.1. Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de He una base cc-pVDZ, a excepción de los valores obtenidos para el octámero y decámero, todos los resultados fueron calculados con PSI4.	56
7.2. Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de Li con una base cc-pVDZ. A excepción de los valores obtenidos para el tetrámero, todos los resultados fueron calculados con PSI4.	57
7.3. Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de Be con una base cc-pVDZ. Los valores obtenidos para el monómero y el dímero fueron calculados con PSI4.	57
7.4. Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de B con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.	57
7.5. Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de C con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.	57
7.6. Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de N con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.	58
7.7. Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de F con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.	58

Abreviaciones

CA	Cúmulos A coplados.
CASD	Cúmulos A coplados con excitaciones S imples y D obles.
CASDT	Cúmulos A coplados con excitaciones S imples, D obles y T riples.
CASDTQ	Cúmulos A coplados con excitaciones S imples, D obles, T riples y
CBC	Conjunto de B ase C ompleto.
ETB	E rror de T runcamiento de la B ase.
FSNE	F unciones de S imetría de N E lectrones.
HF	H artree F ock
ICAC	I nteracción de C onfiguraciones A ltamente C orrelacionado.
ICC	I nteracción de C onfiguraciones C ompleto.
ICS	I nteracción de C onfiguraciones S eleccionado.
ICSD	I nteracción de C onfiguraciones con excitaciones S imples y D obles.
ICSDT	I nteracción de C onfiguraciones con excitaciones S imples, D obles y T riples.
ICSDTQ	I nteracción de C onfiguraciones con excitaciones S imples, D obles, T riples y
	C úadruples.
ICSDTQQn	I nteracción de C onfiguraciones con excitaciones S imples, D obles, T riples,
	C úadruples y Q úintuples
ICSDTQQnSx	I nteracción de C onfiguraciones con excitaciones S imples, D obles, T riples,
	C úadruples, Q úintuples y S éxtuples.
ICSTE	I nteracción de C onfiguraciones S eleccionado con E rror de T runcamiento y
	A nálisis de S ensibilidad.
ICSP	I nteracción de C onfiguraciones S eleccionado y P or P artes.
RADZ	R oos A ugmented D ouble Z eta.
RHF	R estricted H artree F ock.
SCF	S elf C onsistent F ield.

TP **T**eoría de **P**erturbaciones.

Capítulo 1

Introducción

Esta tesis contiene los resultados de la primera implementación del método de interacción de configuraciones con error de truncamiento y análisis de sensibilidad (ICSET) [1] y de interacción de configuraciones seleccionado y por partes (ICSPP) [2] aplicados a sistemas moleculares. Estos métodos tienen como objetivo la obtención de espacios modelo que permitan el cálculo de cotas superiores a la energía que aproximen los valores propios obtenidos con los métodos de interacción de configuraciones (IC) convencionales.

En conjunto, ambos métodos permiten:

- a) la construcción de un espacio modelo que contenga las configuraciones cuantitativamente más importantes con respecto al espacio completo de interacción de configuraciones.
- b) la evaluación cuantitativa de los errores de truncamiento de dicha selección.

Lo anterior hace posible obtener soluciones numéricas de la ecuación de Schrödinger con una exactitud predeterminada, lo que es uno de los objetivos de una teoría de estructura electrónica[4].

Éstas nuevas metodologías (que en conjunto y por simplicidad denominaremos simplemente IC seleccionado (ICS)) presentan varias ventajas con respecto a otras maneras de llevar a cabo cálculos IC [5]. Las técnicas de ICS permiten realizar muy buenas aproximaciones a cálculos de IC completo (ICC) [1, 2] e incluyen nuevas posibilidades inexistentes en otros métodos de estructura electrónica, como son las cotas de error internas (sin recurrir a los datos experimentales) y los análisis de sensibilidad [1].

El uso de ICS ya ha probado su desempeño en el estudio de sistemas atómicos donde, entre otras cosas, han permitido obtener la energía del emblemático estado fundamental del Ne con un error 30 veces menor que el de la competencia [1]. Sin embargo, queda por explorar su uso en el estudio de sistemas moleculares.

El método de IC [5] proporciona la manera más general para el tratamiento de la correlación electrónica [6]: estados excitados, disociación molecular, resonancias y probabilidades de transición.

IC dominó los primeros 50 años de la química cuántica [7] antes de que ser reemplazado por los métodos basados en la teoría de perturbaciones de muchos cuerpos y por los métodos de cúmulos acoplados, sobre todo en el estudio de estados fundamentales y geometrías de equilibrio. La transición empezó cuando se señaló la falta de consistencia en tamaño y extensión en tamaño que presentan la mayoría de los métodos de IC [7–9].

A pesar de la falta de consistencia en tamaño y extensión en tamaño, se considera que los métodos de IC altamente correlacionados (ICAC) (aquellos que incluyen excitaciones más allá de simples y dobles) pueden proporcionar superficies de energía potencial muy confiables, describir de manera adecuada estados electrónicos excitados y producir valores de referencia de la energías y propiedades de sistemas moleculares para su uso en la calibración de los métodos más aproximados [5, 10–12].

En su formas más sofisticadas los métodos ICAC son capaces de reportar valores con una exactitud mayor que la de los mejores resultados experimentales [13]. Sin embargo, para llegar a ese nivel se deben tomar en cuenta las correcciones relativistas, de electrodinámica cuántica, a la aproximación de Born-Oppenheimer, entre otras [14].

El cálculo de propiedades moleculares tan exactas es necesario cuando se quiere probar fenómenos tan extravagantes como la existencia de agua en la superficie solar o bien para mejorar los modelos actuales del efecto invernadero en la tierra [14, 15].

El método de ICS debería permitir obtener resultados IC menos con la misma calidad que los reportados por Polyansky *et al.* [14, 15] cuando se necesite de gran exactitud, pero es lo suficientemente flexible para poder producir valores de energía de sistemas moleculares de mayor tamaño con una nivel de exactitud menor que puede ser controlada por el usuario.

Al respecto, uno de los desafíos abiertos de la química cuántica es la descripción exacta de los sistemas cercanos a la degeneración, lo que sucede cuando se rompe un enlace químico o en la descripción de la estructura electrónica de los elementos de la primera serie de transición.

Al comenzar mi trabajo de doctorado, los estudios de ICS que se habían podido llevar a cabo estaban limitados a sistemas atómicos (los cuales no dejan de ser estudios relevantes, ya que permiten comprender de mejor manera el problema de la correlación electrónica y representan un campo donde existe una amplia tarea por realizar) como parte de mi trabajo de doctorado me encargué de extender metodologías, algoritmos y programas de cómputo desarrollados en nuestro grupo de investigación (ATMOL y AUTOCL) para poder estudiar moléculas pequeñas, esencialmente moléculas diatómicas y triatómicas. Con estos programas hemos podido tratar sistemas de prueba con hasta 20 electrones en 50 orbitales (10 átomos de Helio con una base cc-pVDZ) .

ATMOL (160,000 líneas de código) es un programa de cálculo de estructura electrónica cuyo propósito es la determinación de funciones de onda, valores de expectación y transiciones radiativas. AUTOCL (100,000 líneas de código) es un programa cuya finalidad es la producción de listas de configuraciones de manera general y automática. Ambos permiten realizar tanto cálculos IC no relativistas como relativistas y han sido desarrollados y mantenidos por el Dr. Carlos Bunge y colaboradores durante los últimos 50 años. En particular, AUTOCL fue escrito originalmente en 1966 por la Dra. Annik Vivier-Bunge (actualmente en la UAM-Iztapalapa).

La implementación de ICS para el ámbito molecular permite obtener energías electrónicas con una exactitud del orden de los 160 microHartree (0.1 kcal/mol) o mejor. En general, éste nivel de exactitud se considera más que suficiente para llevar a cabo casi cualquier estudio que se encuentre dentro del ámbito de la química cuántica. Para probar la reciente implementación de ICS reproducimos valores de referencia de la energía para el estado fundamental de H₂O.

Como primer paso reproducimos los valores de las energías de IC reportados por Olsen *et al.* [3], para una distancia H-O igual a la distancia de equilibrio (R_{eq}) y hasta 8 R_{eq} , utilizando una base cc-pVDZ (24 orbitales). Posteriormente, reproducimos las estimaciones del ICC de GKL-Chan y M. Head-Gordon [16] para R_{eq} en una base triple zeta (Roos Augmented Double Zeta Atomic Natural Orbitals)(41 orbitales) [17].

Una vez llevada a cabo la validación de nuestra implementación decidimos proseguir nuestros estudios de la molécula de H_2O en estado de equilibrio. Para realizar esta tarea decidimos hacer uso del conjunto de funciones de base cc-pCVnZ con $n = \text{D, T, Q, 5}$ y 6 con el fin de poder llevar a cabo la extrapolación a base infinita con el conjunto de resultados.

Logramos obtener la mejor cota superior reportada en la literatura a la energía del estado fundamental de la molécula de agua con un ICS a nivel de ICSDTQQnSx, usando una base cc-pCV6Z. El valor calculado (76.4343 Hartrees) [18] recupera más del 98.8 por ciento de la energía de correlación del sistema y se encuentra arriba del valor “experimental” por menos de 3 kcal/mol [10].

Se llevó a cabo un análisis de la convergencia de la contribución a la energía de correlación por orden de excitación y se encontró que (en este caso) para recuperar a nivel de 0.01 miliHartree la contribución a la energía de las excitaciones quintuples y séxtuples no es necesario utilizar una de mayor a cc-pCVQZ.

Posteriormente, con el fin de evaluar el desempeño de diferentes ICAC al representar el proceso de rompimiento simultáneo de los enlaces O-H en la molécula de H_2O llevamos a cabo una revisión de los datos obtenidos por Olsen *et al.*, a distintas distancias de equilibrio, y reproducimos todos los valores de IC que ellos reportan utilizando una función de onda de RHF como referencia.

Al llevar a cabo el estudio de la disociación simétrica de la molécula de agua, se descubrió una manera de optimizar las funciones de onda de referencia usadas en el desarrollo de la función de onda de ICS. Esas funciones de onda son los orbitales naturales obtenidos de la diagonalización sucesiva las matrices densidad de ICAC (interacción de configuraciones con excitaciones simples, dobles y triples (ICSDT) o interacción de configuraciones con excitaciones simples, dobles, triples y cuádruples (ICSDTC). El resultado de dicho proceso es la optimización de los orbitales naturales que conlleva a una función de onda de ICS mucho más compacta y que converge con mayor rapidez a los valores de ICC, respecto al uso de orbitales canónicos de HF o los orbitales naturales de un interacción de configuraciones con excitaciones simple y dobles (ICSD).

Los resultados de ICS obtenidos con este tipo de orbitales naturales optimizados presentan cada vez un menor error de consistencia en tamaño, y en el caso de ICS hasta séxtuples los valores de la energía obtenidos convergen a los de ICC a nivel de 1 microHartree. Posteriormente, se llevo a cabo un conjunto de cálculos de ICSDTQQnSx (llevados a cabo con el programa PSI4 [19]) para todas las diatómicas del segundo período (a excepción del O₂) y se encontró que el mayor error de consistencia en tamaño se presento en el caso del N₂ y no en el caso de F₂ como era de esperarse. Finalmente, se llevó a cabo un estudio similar para distintos conjunto de átomos idénticos (He, Li y Be) separados por una distancia de al menos 10 Å y en cada uno de los casos se encontró que las energías calculadas de ICS hasta séxtuples son múltiplos exactos (con errores menores a 10 microHartree) de las correspondientes energías para los átomos aislados. De estos últimos sistemas los siguientes oligómeros: He₁₀, Li₄, Be₃ representan los sistemas estudiados más grandes.

Capítulo 2

Antecedentes generales de Interacción de Configuraciones

El propósito de éste capítulo es esbozar de manera general, los fundamentos de los métodos IC utilizados en éste trabajo.

Dichos métodos fueron severamente criticados desde hace más de 30 años [7, 9]. Sin embargo, el interés en este tipo de métodos reside principalmente en la necesidad de valores de referencia confiables y en la facilidad de desarrollar métodos multirreferenciales basados en ellos[7] los cuales continúan siendo los métodos de elección en distintas situaciones.

2.1. Ecuación de Schrödinger independiente del tiempo para moléculas.

En principio, la estructura electrónica y las propiedades de cualquier molécula, en cualquiera de sus estados estacionarios, pueden ser determinadas resolviendo la ecuación de Schrödinger (independiente del tiempo) [20]. Para un sistema de n electrones moviéndose en el campo de potencial debido a un conjunto de núcleos, esta ecuación toma la siguiente forma:

$$\hat{H}\Psi(x_1, x_2, \dots, x_n) = E\Psi(x_1, x_2, \dots, x_n), \quad (2.1)$$

donde \hat{H} es el operador Hamiltoniano:

$$\hat{H} = \sum_i h(i) + \frac{1}{2} \sum_{i,j} \frac{e^2}{4\pi\epsilon_0 r_{ij}}, \quad (2.2)$$

El segundo término de la expresión corresponde a la repulsión electrostática entre todos los pares de electrones ij , y el primer término representa el operador Hamiltoniano para el i -ésimo electrón:

$$h(i) = -\frac{\hbar^2}{2m_e} \nabla^2(i) - \sum_n \frac{Z_n e^2}{4\pi\epsilon_0 r_{ni}}, \quad (2.3)$$

Donde el primer término corresponde al operador energía cinética y el segundo término la energía potencial del i -ésimo electrón en campo creado por todos los núcleos del sistema.

2.2. Interacción de Configuraciones

El método IC es reconocido como el método más general para tratar la correlación electrónica [6]. Tiene como origen la aplicación del método de Ritz [21] en el cálculo de funciones de onda electrónicas. El método IC ha sido utilizado en cálculos atómicos o moleculares, ya sea en términos de orbitales atómicos, orbitales moleculares e incluso dentro del método de enlace valencia.

El IC completo (ICC), Ψ^{ICC} , se basa en una función de onda, que se expresa como combinación lineal de determinantes de Slater [22] (D_0 , D_i^a , $D_{i,j}^{a,b}$, $D_{i,j,k}^{a,b,c}$, etc.), cuyos coeficientes $c_{i,j,k,\dots}^{a,b,c,\dots}$ se determinan de manera variacional:

$$\Psi^{ICC} = c_0 D_0 + \sum_{a,i} c_i^a D_i^a + \sum_{\substack{a<b \\ i<j}} c_{ij}^{ab} D_{ij}^{ab} + \sum_{\substack{a<b<c \\ i<j<k}} c_{ijk}^{abc} D_{ijk}^{abc} + \dots, \quad (2.4)$$

Donde D_0 , en general, es un determinante de Hartree-Fock (HF) denominado determinante de referencia. Los términos de las sumas representan determinantes en los cuales se han sustituido los espín-orbitales ocupados (i, j, k, \dots) en D_0 por espín-orbitales que son denominados espín-orbitales de correlación (a, b, c, \dots). D_i^a , $D_{i,j}^{a,b}$, $D_{i,j,k}^{a,b,c}$, etc., se denominan, respectivamente, excitaciones simples, dobles, triples, etc., hasta n -tuples, donde n es el número de electrones.

Los espín-orbitales $\{ \phi_{ki}; i = 1, 2, \dots, N \}$ se expresan de la manera tradicional como productos de orbitales χ_i con funciones de espín α o β , con eigenvalores $+\frac{1}{2}$ y $-\frac{1}{2}$ de la proyección del momento angular de espín, respectivamente. El truncamiento a diferentes niveles de excitación da, sucesivamente, IC con excitaciones simples (IC-S), IC con simples y dobles (ICSD), IC con simples, dobles y triples (ICSdT), IC hasta cuádruples (ICSdTC), etc.

La función de onda completa tendría que contener hasta determinantes n -excitados, representando la solución exacta para una base dada de espín-orbitales y dentro del régimen no relativista [20]. El número de términos en (2.4) crece como el número de espín-orbitales de referencia ocupados elevado al nivel de excitación multiplicado por el número de espín-orbitales de correlación también elevado al nivel de excitación, esto es, el crecimiento es exponencial.

El desarrollo (2.4) en términos de determinantes de Slater puede compactarse drásticamente si se toma en cuenta la simetría del sistema [23]. Entonces puede reescribirse como:

$$\Psi^{ICC} = \sum_{K=1}^{K_x} \sum_{g=1}^{g_K} F_{gK} C_{gK}, \quad (2.5)$$

donde F_{gK} designa a las llamadas funciones (ortonormales) de simetría de n electrones (FSNE). C_{gK} representa los coeficientes de IC, K y g constituyen las etiquetas de las configuraciones y la degeneración respectivamente, mientras K_x y g_k son los límites superiores de las sumas respectivas.

Al utilizar como función de onda la expresión (2.5), las soluciones de la ecuación de Schrödinger pueden aproximarse por:

$$\mathbf{H}\mathbf{C}_n = E_n^{ICC}\mathbf{C}_n, \quad (2.6)$$

donde \mathbf{H} es la representación del Hamiltoniano en una base ortonormal de FSNE, las cuales son ser expresadas como combinaciones lineales de determinantes de Slater. E_n^{ICC} es el eigenvalor correspondiente al eigenvector \mathbf{C}_n .

2.3. Funciones de simetría de n electrones.

Las FSNE son funciones propias de operadores que conmutan con el Hamiltoniano [23]. En el caso de átomos las FSNE son funciones propias de los operadores de momento angular de espín \mathbf{S}_z y \mathbf{S}^2 , así como de los operadores de momento angular orbital \mathbf{L}_z y \mathbf{L}^2 , mientras que en el caso de moléculas, generalmente se utilizan FSNE de simetría molecular y de espín.

En los casos más simples las FSNE pueden construirse por inspección o por “ensayo y error”. Para sistemas un poco más complicados una opción puede ser llevar a cabo la diagonalización de la matriz correspondiente a cada uno de los operadores de simetría pertinente [24], pero para cálculos de gran escala, la construcción de estas funciones debe de llevarse a cabo de manera sistemática [23].

Aunque hay varias maneras de construir sistemáticamente las FSNE, en el presente trabajo su construcción se llevó a cabo de acuerdo al formalismo introducido por Willis [25] y elaborado por Löwdin [26–29] usando operadores de proyección, mediante el programa AUTOCL.

Podemos expresar una FSNE como una combinación lineal de n_K determinantes de Slater, D_{iK} , a la cual se le aplica el correspondiente operador de proyección $O(M, k)$:

$$F_{gK} = O(M, k) \sum_{i=1}^g D_{iK} b_i^g = \sum_{i=1}^{n_K} D_{iK} c_i^g, \quad g = 1, \dots, g_K, \quad (2.7)$$

Al llevar a cabo éste proceso se produce la correspondiente eigenfunción del operador de momento angular \mathbf{M}^2 pertinente, F_{gK} . La función F_{gK} corresponde a las representaciones irreducibles \mathbf{k} etiquetadas por los números cuánticos M y k , de los operadores \mathbf{M}^2 y \mathbf{M}_z , respectivamente. El índice g representa la etiqueta perteneciente a los elementos degenerados de la configuración K . $O(M, k)$ viene dado por:

$$O(M, k) = (2M + 1)! \frac{(M + k)!}{(M - k)!} \sum_{\nu=1}^{M_X - M} \frac{(-1)^\nu \mathbf{M}_-^{M-k+\nu} \mathbf{M}_+^{M-k+\nu}}{\nu!(2M + \nu + 1)!}, \quad (2.8)$$

La implementación con la que cuenta ATMOL permite producir FSNE con bases de orbitales de momento angular elevado, hasta $l = 30$.

2.4. Orbitales naturales.

La matriz densidad reducida de primer orden $D(x : x')$ se define como [30]:

$$D(x : x') = N \int \Psi(x, x_2, \dots, x_n) \Psi^*(x', x_2, \dots, x_n) dx_2 \dots dx_n. \quad (2.9)$$

En términos de orbitales X_k , $D(x : x')$ puede expresarse como:

$$D(x : x') = \sum_{i,j} d_{ij} X_i(x) X_j^*(x'). \quad (2.10)$$

Si se define un operador integral \hat{D} :

$$\hat{D}f = \int D(x : x') f(x') dx, \quad (2.11)$$

cuyas funciones propias φ_i , se denominan orbitales naturales. Dichas eigenfunciones vienen dadas por:

$$\hat{D}\varphi_i = n_i \varphi_i, \quad (2.12)$$

donde los correspondientes valores propios n_i se denominan números de ocupación.

Los orbitales naturales presentan la característica de reducir a su forma diagonal a la matriz definida en (2.10).

$$D(x : x') = \sum_i n_i \varphi_i(x) \varphi_i^*(x'). \quad (2.13)$$

La importancia de los orbitales naturales en los cálculos IC es que el uso de los números de ocupación permite un ordenamiento jerárquico de la base de orbitales facilitando tanto la identificación de configuraciones importantes en el desarrollo de la función de onda de IC como haciendo posible un análisis de sensibilidad [1].

2.5. Error de truncamiento de la base.

A causa de que las bases de orbitales son finitas, las E_n de la ecuación (2.1) difieren de las E_n^{ICC} en una cantidad que se denomina error de truncamiento de la base (ETB) [1], ΔE_n^{ETB} .

$$E_n = E_n^{ICC} + \Delta E_n^{ETB} \quad (2.14)$$

En el caso de átomos, todos los métodos para obtener ΔE_n^{ETB} están basados en el uso de orbitales naturales [31, 32]. Para el caso de las moléculas, la mayoría de los métodos se basan en el uso de los conjuntos de base especialmente diseñados para converger sistemáticamente a E_n [33].

2.6. Interacción de Configuraciones frente a otros métodos de primeros principios

El principal problema que presenta ICC es el aumento exponencial en el número de configuraciones para órdenes de excitación sucesivos [24]. Por lo tanto, ICC es computacionalmente inviable excepto para sistemas pequeños y mientras se utilicen bases de orbitales no demasiado grandes [3, 6, 9].

Consecuentemente, el desafío de IC es desarrollar aproximaciones a ICC para bases de orbitales de tamaño adecuado, las cuales presenten un ETB que esté acotado dentro de límites que puedan llegar a conocerse, y para sistemas que requieran una exactitud mayor a la que se puede tener con otros métodos establecidos.

Hace poco más de 30 años el método IC sufrió un golpe casi decisivo, al ponerse en evidencia que la energía de IC truncado a cualquier orden de excitación no es consistente en tamaño [7, 9], esto es, la energía de dos sistemas no interactuantes A y B no es igual a la suma de las energías de A y B calculadas por separado. Por el contrario, HF, teoría de perturbaciones (TP) a cualquier orden, y cúmulos acoplados (CA), son consistentes en tamaño. A pesar de esto, otros métodos como IC-multirreferencial [34], que tampoco es consistente en tamaño, continuó desarrollándose y compitiendo con ventaja con CA en el estudio de sistemas intrínsecamente multirreferenciales, o cual es muy frecuente en el cálculo de superficies de energía electrónica.

Sin embargo, a pesar de la “mancha” que representa la carencia de consistencia en tamaño, IC es competitivo aún con CA, la estrella de los métodos consistentes en tamaño: estudios llevados a cabo sobre la molécula de agua [3, 35] muestran que los resultados de energía de CASDT (CA que incluye excitaciones hasta tercer orden), recuperan un porcentaje de la energía de correlación similar a los obtenidos con un tratamiento de ICSDTC. Esto se debe al carácter variacional de las energías de IC, que no tienen las de CA.

Capítulo 3

Interacción de Configuraciones seleccionado.

3.1. Antecedentes.

Para poder realizar un cálculo IC de manera práctica, la simplificación más evidente que podemos llevar a cabo es limitar el número de configuraciones que incluimos en nuestra función de onda [5]. Lo cual se lleva a cabo siguiendo una serie de criterios, de los cuales algunos de los más comunes se mencionarán a continuación.

3.1.1. Truncamiento por nivel de excitación.

La manera más “natural” de truncar la función de onda es mediante la supresión de aquellos términos que corresponden a las excitaciones de mayor orden [5]. La forma más común de implementar esto es incluir en la función de onda únicamente excitaciones simples y dobles (ICSD). El principal motivo para realizar esto es que un cálculo de ICSD, presenta un costo computacional aceptable.

Otra justificación para llevar a cabo esta selección es que en un modelo con un único determinante o función de referencia los únicos términos del desarrollo de la función de onda de IC que interactúan directamente con la referencia son aquellos que corresponden a excitaciones simples y dobles; esto se debe a que el Hamiltoniano contiene únicamente términos que representan la interacción entre dos cuerpos (electrones y electrón-núcleo, en nuestro caso).

Un cálculo ICSD de un sistema molecular cerca de la geometría de equilibrio puede recuperar hasta 95 % de la energía de correlación [5, 6], sin embargo, para estudiar procesos de disociación molecular el tratamiento de la correlación electrónica a nivel de ICSD es totalmente insuficiente en casi cualquiera circunstancia (siendo la excepción más notable el caso de la disociación de la molécula de hidrógeno).

El siguiente nivel de aproximación más común lo representa ICSDTC, que sólo es posible llevar a cabo de manera convencional para sistemas de tamaño mediano, con no más de 10-20 electrones, utilizando bases relativamente pequeñas, y a un costo prohibitivo [5].

3.1.2. Selección de determinantes bajo criterios de número de ocupación.

Otra manera de seleccionar FSNE se basa en criterios de números de ocupación de los orbitales virtuales. Al tomar éste criterio, las FSNE que representen excitaciones de orbitales de la referencia hacia orbitales virtuales cuyo número de ocupación sea menor a cierto umbral no son incluidas en Ψ .

También se puede utilizar un criterio de selección que “congele” a los orbitales ocupados de menor energía, esto es, los del corión[5], con el argumento de que sus excitaciones contribuirán muy poco a la descripción del sistema, de modo tal que podemos ahora hablar de orbitales “activos”, entendiendo por éstos únicamente aquellos que se toman en cuenta para formar las excitaciones de las FSNE. El conjunto de los orbitales que permanecen “activos” tanto ocupados como virtuales constituye el denominado “espacio activo” [5].

Utilizando estos dos criterios podemos reducir de manera considerable el número de FSNE incluidas en Ψ . En éste trabajo, el espacio completo de orbitales coincide con el espacio activo. Sin embargo, tanto ATMOL como AUTOCL ya son capaces de llevar a cabo cálculos de IC utilizando este tipo de aproximaciones.

3.1.3. Selección bajo criterios de espacios interactuantes.

Frecuentemente, en sistemas de capas abiertas el número de términos que se incorporan en la función de onda de ICSD es demasiado grande, por lo cual debe ser seleccionado. Para ello puede hacerse uso del concepto de espacios interactuantes [36].

La idea detrás de esto es conceptualmente simple: el espacio de orden cero viene dado por Ψ_0 y espacio de orden uno por combinaciones lineales de F_{gK} que representen excitaciones simple y dobles (Ψ_1 y Ψ_2) con elementos de matriz no nulos con respecto a Ψ_0 , $\langle \Psi_0 | H | a\Psi_1 + b\Psi_2 \rangle \neq 0$.

Lo útil de éste concepto es que el espacio interactuante de primer orden puede llegar a reducirse en un orden de magnitud con respecto al espacio completo de excitaciones simples y dobles sin afectar la energía de manera significativa. Nótese que las FSNE con orden orbital de excitación uno y dos están representadas por determinantes de orden de excitación uno, dos y superiores. Estos últimos, que constituyen la gran mayoría, son los que quedan fuera del espacio interactuante de orden uno. De manera que el espacio interactuante de orden uno es el espacio de FSNE que se obtiene como combinaciones lineales de determinantes de Slater con grado de excitación de orden uno y dos [37].

Sin embargo, es imposible estimar *a priori* el efecto de truncamiento en la energía del complemento del espacio interactuante de orden uno, por lo cual, su utilidad práctica se limita al uso de funcionales compactas para optimizar los parámetros no lineales de la base de primitivas en sistemas atómicos.

3.1.4. Selección bajo criterios de energía y de coeficientes de IC.

Hasta aquí hemos mencionado maneras de seleccionar términos de la función de IC, que aunque resultan muy útiles, tienen el principal defecto de que no permiten calcular de manera directa el error de truncamiento en la energía debido a los términos descartados.

La idea de seleccionar las FSNE que se incorporan en la función de onda con base a la estimación de su contribución a la energía ha sido desarrollada durante varios años [5, 24], dando como resultado toda una serie de técnicas que se incorporan de manera habitual en casi cualquier cálculo de IC.

La mayoría de estos métodos están basados en la teoría de perturbaciones de Rayleigh-Schrödinger, mediante la cual podemos, por ejemplo, calcular la contribución a la energía de segundo orden:

$$\Delta E_i = \frac{|\langle \Psi_i | H | \Psi_0 \rangle|^2}{\langle \Psi_0 | H | \Psi_0 \rangle - \langle \Psi_1 | H | \Psi_1 \rangle}. \quad (3.1)$$

Si dicha contribución a la energía ΔE_i es menor que un cierto umbral podemos descartar la configuración Ψ_i , mientras que ΔE_i se acumula para proporcionar el error de truncamiento total ΔE .

Además de la contribución a la energía podemos calcular los coeficientes de expansión de cada configuración, C_i :

$$C_i = \frac{\langle \Psi_i | H | \Psi_0 \rangle}{\langle \Psi_0 | H | \Psi_0 \rangle - \langle \Psi_1 | H | \Psi_1 \rangle}, \quad (3.2)$$

los cuales también podemos usar en un criterio de selección de configuraciones. De manera tal que incluyamos únicamente aquellas configuraciones cuyo valor de $|C_i|$ sea mayor a cierto umbral.

Dentro del esquema de metodologías de selección basadas en contribuciones a la energía calculadas *a priori* a partir de la teoría de perturbaciones de Rayleigh-Schrödinger, la primera de ellas en representar una manera eficiente y sistemática de seleccionar la constituyó el método denominado Interacción de configuraciones perturbando una función de onda multiconfiguracional seleccionada de manera iterativa, mejor conocido como CIPSI [38, 39]. CIPSI goza de gran popularidad y cabe mencionar representó la manera de llevar a cabo la primeros cálculos moleculares post-HF en América Latina.

Ya que en éste caso Ψ_0 en (3.1) es multiconfiguracional con excitaciones simples y dobles, el error de truncamiento ΔE contiene el efecto de truncamiento de algunas excitaciones triples y cuádruples, pero no de todas ellas, y mucho menos de excitaciones quintuples y superiores.

El principal problema que presenta éste tipo de selección es el número de términos a evaluar. En el peor de los casos tenemos que calcular cada una de las FSNE que podrían ser incluidas en la función de onda, y recordemos que el número de ellas crece rápidamente con respecto al orden de excitación de las configuraciones.

3.2. Interacción de Configuraciones con errores de truncamiento.

Como hemos visto, mucho del trabajo realizado en el intento de acercarse de manera sistemática a los resultados de ICC tiene que ver con la utilización criterios de selección. Sin embargo, el uso de estos criterios no aborda de manera adecuada el cálculo del consiguiente error de truncamiento.

En los últimos años, se ha desarrollado el método denominado interacción de configuraciones seleccionado con errores de truncamiento [1, 2]. Éste método proporciona una manera práctica de llevar a cabo cálculos de IC-Seleccionado que aproximan de excelente manera las energías de ICC de sistemas de pocos electrones, digamos 10 electrones, con bases muy grandes y con la utilización de recursos de cómputo y tiempo de investigador razonables.

Las principales características de la metodología de interacción de configuraciones seleccionado con errores de truncamiento (ICS) son:

- (1) los errores de truncamiento de la energía, son evaluados *a priori* mediante el uso de fórmulas que han probado ser eficaces y cuantitativamente correctas.
- (2) la selección se lleva a cabo sobre el conjunto de todas las FSNE asociadas a una configuración en vez de hacerlo para cada una de las FSNE.

Debido a lo anterior, la selección aproxima el ICC de manera sistemática y confiable. A continuación se presentan los detalles más importantes de dicha metodología.

3.2.1. La fórmula de Brown.

Retomando el desarrollo de la función de onda de IC en términos de FSNE:

$$\Psi^{ICC} = \sum_{K=1}^{K_x} \sum_{g=1}^{g_K} F_{gK} C_{gK}, \quad (2.5)$$

Comenzamos por obtener una expresión para la contribución a la energía ΔE_{gK} de cada una de las F_{gK} . Primero suponemos que la función de onda a la cual le hemos retirado una determinada F_{gK} puede ser expresada como $\Psi(-F_{gK}) = N(\Psi - F_{gK})$, donde N representa un factor de normalización. De manera tal que la función de onda permaneció igual que la original (con todos los coeficientes C_{gK} manteniendo el mismo valor que en Ψ), excepto por el factor de renormalización N .

ΔE_{gK} puede expresarse como:

$$\Delta E_{gK} = \langle \Psi | H | \Psi \rangle - \langle \Psi(-F_{gK}) | H | \Psi(-F_{gK}) \rangle. \quad (3.3)$$

A partir de la cual se obtiene la expresión:

$$\Delta E_{gK} = \frac{(\langle \Psi | H | \Psi \rangle - H_{gK,gK})^2 C_{gK}^2}{1 - C_{gK}^2}, \quad (3.4)$$

o bien:

$$\Delta E_{gK} = \frac{(E - H_{gK,gK})^2 C_{gK}^2}{1 - C_{gK}^2}, \quad (3.5)$$

que se conoce con el nombre de fórmula de Brown [40]. Ésta fórmula nos permita calcular la contribución a la energía ΔE_{gK} de la configuración F_{gK} , en términos de su coeficiente de expansión C_{gK} , el elemento diagonal $H_{gK,gK}$ y un valor aproximado de la energía del sistema, E . Cabe mencionar que el valor utilizado para representar E no necesita ser demasiado exacto, ya que $(E - H_{gK,gK})^2$ es, generalmente, del orden de las decenas o hasta miles de Hartree.

La aproximación será mejor conforme menor sea la contribución a la energía de una determinada configuración. Por lo cual, dicha fórmula proporciona un criterio de selección excelente, pues son precisamente aquellas configuraciones que contribuyan en menor medida a la energía con las que mejor funciona la fórmula de Brown. Así, podemos utilizarla como criterio de selección de configuraciones de excitaciones triples, cuádruples, quintuples, etc. y el resultado obtenido será cada vez mejor. Al mismo tiempo, como bono extra obtenemos un error de truncamiento cuantitativo para los términos descartados.

Por supuesto, para conocer el valor de ΔE_{gK} , tenemos que conocer el valor de cada uno de los términos de la expresión (3.5). Para E podemos utilizar el valor de un cálculo de ICSD, tomando en cuenta que dicho valor representa alrededor del 95 % de la energía total del sistema. El elemento diagonal $H_{gK,gK}$ puede evaluarse convencionalmente. Por lo tanto, el único elemento faltante es el valor de los coeficientes C_{gK} de cada configuración.

Sin embargo, la obtención de los valores de C_{gK} representa un serio desafío, puesto que la manera directa de obtener estos coeficientes es precisamente resolver el problema respectivo de valores propios, ecuación (2.6). Calculando los ΔE_{gK} a posteriori se puede ir simplificando el desarrollo de (2.5) pero el costo es muy alto y la mecánica de cálculo se vuelve tediosa [41].

Además, el uso de la fórmula de Brown, expresada en (3.5), no representa una manera eficiente de llevar a cabo la selección. Esto se debe a la dificultad que implica obtener expresiones de C_{gK} para cada uno de los elementos degenerados, g , de una configuración, K . Sin embargo, como se mostrará más adelante, es posible predecir los valores de los coeficientes B_K , correspondientes al desarrollo de función de onda de IC en términos de configuraciones de simetría normalizadas, G_k :

$$\Psi = \sum_{K=1}^{K_x} G_K B_K, \quad (3.6)$$

Donde G_K se define como:

$$G_K = N_K \sum_{g=1}^{g_K} F_{gK} C_{gK}, \quad (3.7)$$

con:

$$B_K = \frac{1}{N_K}, \quad N_K = \sqrt{\frac{1}{\sum_{g=1}^{g_K} C_{gK}^2}}. \quad (3.8)$$

A partir de la cual podemos encontrar la correspondiente expresión de ΔE_K :

$$\Delta E_K = (E - H_{KK})B_K^2/(1 - B_K^2), \quad (3.9)$$

en la cual podemos evaluar E de la manera que se planteó para (3.5). $H_{K,K}$ puede ser aproximado por:

$$H_{KK} \approx \langle D_{iK} | H | D_{iK} \rangle, \quad (3.10)$$

donde D_{iK} es cualquier determinante de K . Al llevar a cabo ésta simplificación se evita el cálculo de las FSNE, permitiendo la evaluación de ΔE_K aun cuando $g_K \gg 1000$ y el número de determinantes $n_K \gg 100000$.

De manera que podemos llevar a cabo un proceso de selección con error de truncamiento de manera mucho más eficiente. Poder llevar a cabo éste proceso es crucial en el caso de sistemas donde se presenta gran degeneración en los términos de la ecuación (2.5), como es el caso de átomos. No obstante, este no es el caso de las moléculas con simetría puntual perteneciente a grupos puntuales abelianos, es importante resaltar que el cálculo de IC aún se llevará a cabo en términos de F_{gK} , en vez de G_K .

3.2.2. Desarrollo de cúmulos y predicción de coeficientes de expansión configuracionales.

Antes de continuar, es importante introducir el concepto de configuraciones conectadas y desconectadas. Para aquellos términos de la expresión (2.5) que corresponden a excitaciones triples, cuádruples, etc. las configuraciones desconectadas son aquellas que pueden expresarse como producto de configuraciones que representan excitaciones de un orden de excitación menor. El resto de las configuraciones se denominan configuraciones conectadas. En moléculas con simetría perteneciente a grupos abelianos, la mayoría de los términos del desarrollo de la función de onda de IC corresponde a configuraciones desconectadas.

Ahora, retomemos la expresión del desarrollo de función de onda de IC en términos de determinantes de Slater:

$$\Psi = D_0 c_0 + \sum_i \sum_a D_i^a c_i^a + \sum_{i < j} \sum_{a < b} D_{ij}^{ab} c_{ij}^{ab} + \sum_{i < j < k} \sum_{a < b < c} D_{ijk}^{abc} c_{ijk}^{abc} + \dots \quad (2.4)$$

En 1962 Oktay Sinanoğlu [42] sugirió que los coeficientes $c_{ijk\dots}^{abc\dots}$, correspondientes a determinantes n-excitados están relacionados de manera aproximada con los coeficientes correspondientes a determinantes de menor grado de excitación.

De manera que, por ejemplo, los coeficientes de excitaciones cuádruples se pueden aproximar a partir de los coeficientes correspondientes a excitaciones dobles:

$$c_{ijkl}^{abcd} \approx c_{ij}^{ab} c_{kl}^{cd} + c_{ik}^{ab} c_{jl}^{cd} + c_{il}^{ab} c_{jk}^{cd}. \quad (3.11)$$

Para encontrar relaciones exactas y generales para evaluar los coeficientes $c_{ijk\dots}^{abc\dots}$, tenemos que apoyarnos en la teoría de cúmulos acoplados (CA). En CA se parte de una función de onda expresada como:

$$\Psi = \exp(T) D_0, \quad (3.12)$$

donde D_0 , es un determinante de referencia, generalmente de HF. T es el operador de cúmulos [43]:

$$T = T_1 + T_2 + T_3 + \dots \quad (3.13)$$

En (3.13) cada uno de los términos es representado en términos de coeficientes $t_{ijk\dots}^{abc\dots}$, llamados amplitudes, y de operadores de creación \hat{a}_a^\dagger y destrucción \hat{a}_i :

$$T_1 = \sum_i \sum_a t_i^a \hat{a}_a^\dagger \hat{a}_i, \quad (3.14)$$

$$T_2 = \sum_{i < j} \sum_{a < b} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i, \quad (3.15)$$

$$T_3 = \sum_{i < j < k} \sum_{a < b < c} t_{ijk}^{abc} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_c^\dagger \hat{a}_k \hat{a}_j \hat{a}_i, \quad (3.16)$$

Si desarrollamos la exponencial en (3.12) y agrupamos términos, podemos obtener relaciones exactas que nos permiten relacionar los coeficientes del IC de la expresión (2.4), $c_{ijk\dots}^{abc\dots}$, con las amplitudes de cúmulos acoplados, $t_{ijk\dots}^{abc\dots}$.

$$\frac{1}{c_0} c_i^a = t_i^a, \quad (3.17)$$

$$\frac{1}{c_0} c_{ij}^{ab} = t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a, \quad (3.18)$$

$$\begin{aligned} \frac{1}{c_0} c_{ijk}^{abc} = & t_{ijk}^{abc} + t_i^a t_j^b t_k^c - t_i^b t_j^a t_k^c + t_i^c t_j^a t_k^b - t_i^a t_j^c t_k^b + t_j^b t_i^a t_k^c - t_j^c t_i^b t_k^a + \\ & t_k^a t_i^b t_j^c - t_k^b t_i^c t_j^a + t_k^c t_i^a t_j^b + t_i^a t_j^b t_k^c - t_i^b t_j^c t_k^a - t_i^c t_j^a t_k^b + t_j^b t_i^c t_k^a + t_j^c t_i^a t_k^b - t_i^c t_j^b t_k^a, \end{aligned} \quad (3.19)$$

Mediante una serie de despejes podemos encontrar fórmulas del tipo:

$$\frac{1}{c_0} c_{ij}^{ab} = t_{ij}^{ab} + \frac{1}{c_0^2} (c_i^a c_j^b - c_i^b c_j^a). \quad (3.20)$$

$$\begin{aligned} \frac{1}{c_0} c_{ijk}^{abc} = & t_{ijk}^{abc} + \frac{1}{c_0^2} (c_i^a c_j^b c_k^c - c_i^b c_j^a c_k^c + c_i^c c_j^a c_k^b - c_j^a c_i^b c_k^c + c_j^b c_i^a c_k^c - c_j^c c_i^b c_k^a + c_k^a c_i^b c_j^c - c_k^b c_i^c c_j^a + c_k^c c_i^a c_j^b) \\ & - \frac{2}{c_0^3} (c_i^a c_j^b c_k^c - c_i^b c_j^c c_k^a - c_i^c c_j^a c_k^b + c_i^b c_j^c c_k^a + c_i^c c_j^a c_k^b - c_i^a c_j^c c_k^b) \end{aligned} \quad (3.21)$$

De manera general:

$$\frac{1}{c_0} c_{ijk\dots}^{abc\dots} = t_{ijk\dots}^{abc\dots} + F(c_{ij\dots}^{ab\dots}). \quad (3.22)$$

En las expresiones del tipo (3.22) el único término desconocido es el que corresponde a la amplitud de cúmulo. Para excitaciones triples y de mayor orden, las amplitudes pueden despreciarse. Lo cual se debe a que, a diferencia de los coeficientes de IC, los valores de dichas amplitudes disminuyen rápidamente con respecto al orden de excitación, por analogía con la teoría de cúmulos para gases imperfectos [42].

Entonces, si expresamos la función de onda (2.5) en términos de configuraciones de simetría normalizadas, definidas en (3.7):

$$\Psi = \Phi_0 B_0 + \sum_i \sum_a \Phi_i^a B_i^a + \sum_{i \leq j} \sum_{a \leq b} \Phi_{ij}^{ab} B_{ij}^{ab} + \sum_{i \leq j \leq k} \sum_{a \leq b \leq c} \Phi_{ijk}^{abc} B_{ijk}^{abc} + \dots \quad (3.23)$$

Debido al hecho de que todos los términos degenerados presentes (2.5), se han contraído a un solo término en (3.23), hemos logrado una contracción muy eficiente. La contracción se vuelve mejor cuanto mayor sea el valor de g para un determinado término en (3.23). Nótese que al pasar de espín-orbitales i, j, \dots, a, b, \dots a orbitales i, j, \dots, a, b, \dots los índices pueden ser iguales entre sí.

Uno de los fundamentos del ICS es el uso de un conjunto fórmulas para predecir el valor de los coeficientes $B_{ijk\dots}^{abc\dots}$ con la finalidad de seleccionar configuraciones mediante la fórmula de Brown (3.9). Las nuevas fórmulas fueron “adivinadas” a partir de las fórmulas correspondientes para calcular los coeficientes $c_{ijk\dots}^{abc\dots}$ [1]. Por ejemplo, para predecir el coeficiente B_{ijk}^{abc} en el que no haya signos de igual y todas las degeneraciones sean igual a uno, se puede hacer uso de las mismas fórmulas que se utilizarían para calcular los coeficientes $c_{ijk\dots}^{abc\dots}$.

En el otro extremo, cuando $i = j = k$, $a = b = c$, para calcular los coeficientes se recurre a (3.21) y la expresión que se obtiene es:

$$\frac{1}{B_0} B_{iii}^{aaa} = \hat{B}_{iii}^{aaa} + \frac{1}{B_0^2} B_i^a B_{ii}^{aa} - \frac{2}{B_0^3} B_i^a B_i^a B_i^a. \quad (3.24)$$

En la ecuación (3.24) la amplitud $t_{ijk\dots}^{abc\dots}$, que corresponde al coeficiente \hat{B}_{iii}^{aaa} irreducible en la notación de Sinanoğlu, será despreciado para poder obtener el correspondiente valor de B_{iii}^{aaa} . Para tres índices ocupados hay un total de cuatro posibilidades: ijk , ijj , ijj , iii . De manera análoga, existen cuatro posibilidades abc , abb , aab y aaa para los índices virtuales, con lo cual resulta un total de 16 combinaciones posibles para 16 fórmulas de los coeficientes de excitaciones triples de (3.23).

En general, para excitaciones de orden q , habrá $2^{q-1} \cdot 2^{q-1} = 2^{2q-2}$ posibilidades. Si $q = 6$, el número de fórmulas es $2^{10} = 1024$, requiriendo un código de aproximadamente 64,000 líneas compactas de Fortran. Obviamente dicho código sólo puede ser generado por un programa de cómputo [2], y $q = 6$ representa el límite actual de los programas ATMOL y AUTOCL, tanto para átomos como para moléculas.

Cuando $q = 8$ se obtienen más de 4 millones de líneas de Fortran. Sin embargo, tanto para la mayoría de las aplicaciones atómicas, como para moléculas con grupos abelianos (no más de dos índices iguales) el número de fórmulas se reduce a proporciones manejables. Aunque su discusión y uso eventual será pospuesto para investigaciones futuras.

De ésta manera, podemos evaluar la fórmula de Brown para el caso de la configuraciones G_K desconectadas, pero no para las configuraciones conectadas. Esto se debe a que para éstas, el único término que aparece en la ecuación (3.22) es el que corresponde al coeficiente irreducible de Sinanoğlu. Por lo tanto, aún tenemos que encontrar un criterio similar al que proporciona la fórmula de Brown para seleccionar las configuraciones conectadas.

3.2.3. Criterio de selección para configuraciones conectadas.

En los inicios de ICS se propuso como criterio de selección de las configuraciones conectadas de orden n , el producto de los números de ocupación $P(q, K)$:

$$P(q, K) = \prod_{i=1}^q n_{K,i} \quad (3.25)$$

con muy buenos resultados [1].

Aunque no se contaba con un criterio que diera cuenta de manera cuantitativa del error de truncamiento de energía de esta selección, un (costoso) posterior análisis de sensibilidad realizado de manera sistemática permitía tener control sobre esta situación y dar cuenta del error de truncamiento de la energía.

Sin embargo, en fecha reciente [44], se propuso utilizar una fórmula basada en la fórmula de Brown que sirviera como criterio de selección semi-cuantitativo en la selección de configuraciones conectadas. Dicha fórmula está dada por:

$$\Delta E_K^{\text{pseudoenergía}} = (E - H_{K,K})P(q, K) \quad (3.26)$$

donde $n_{K,i}$ es el número de ocupación del orbital natural de orden i en la configuración K con orden de excitación q . La cantidad $\Delta E_K^{\text{pseudoenergía}}$, dada por (3.26) tiene unidades de energía, pero su poder predictivo es cuantitativamente limitado. Debido a lo anterior, dicha cantidad se considera una pseudoenergía, la cual forma parte de un nuevo criterio de selección, el criterio de las pseudoenergías.

En la práctica, el criterio de las pseudoenergías proporciona un criterio de selección confiable y mucho más eficiente que el simple criterio sugerido por (3.25) [1, 44].

A lo largo de éste capítulo se expuso la necesidad de llevar a cabo una selección de los términos a incluir en la función de onda de IC. Se plantearon diversas maneras de llevar esto a cabo, poniendo énfasis en la necesidad de contar con métodos eficaces y eficientes que además den como resultado estimaciones cuantitativas de los errores de truncamiento a la energía resultantes de dicha selección. Después, se desarrollaron los fundamentos básicos de la metodología de ICS empleada en éste trabajo. En la cual, mediante el uso de la fórmula de Brown y una modificación de ella, se pueden obtener estimaciones cuantitativas y confiables de los errores de truncamiento debidos al proceso de selección de configuraciones. Además, dicha selección se lleva a cabo sobre conjuntos de configuraciones degeneradas, lo cual hace más eficiente el proceso.

La metodología ICS cumple a cabalidad las condiciones que uno desea tener en un buen esquema de selección. La pesadilla del crecimiento exponencial en la complejidad del IC se pospone para sistemas con un mayor número de electrones. Además ICS hace posible, en el presente, el uso de bases de gran tamaño, que quedan fuera de las posibilidades de los métodos de CA actuales.

Capítulo 4

Interacción de Configuraciones seleccionado y por partes.

El fundamento de ICS es escoger un subespacio, S , de un espacio de configuraciones M que represente, por ejemplo, una función de onda de ICC, que contenga el conjunto de configuraciones que contribuyen de manera más importante a la energía [2]. La elección de dicho conjunto se lleva a cabo mediante el uso de las ecuaciones (3.9) y (3.26).

Las contribuciones individuales a la energía de cada una de las configuraciones descartadas, calculadas mediante la fórmula de Brown o sus modificaciones, se acumulan en un error de truncamiento ΔE , con su consiguiente incertidumbre asociada δE .

Por lo tanto, la energía de M puede ser representada como:

$$E_M = E_S + \Delta E + \delta E, \quad (4.1)$$

donde, el valor de E_S se calcula resolviendo el correspondiente problema de valores propios:

$$H^S C = E_s C \quad (4.2)$$

El ICS muestra que el número de configuraciones que incluyen el espacio de configuraciones seleccionado, S , sigue siendo tan grande que resulta complicado hacer uso de las técnicas habituales de resolver el correspondiente problema de valores propios (entre las cuales la más utilizada la representa el denominado método de Davidson [45, 46] y sus modificaciones). Por tanto, es necesario el uso de una técnica diferente que nos permita llevar a cabo dicha tarea.

La técnica que hace posible evaluar el correspondiente problema de valores propios viene dado por el método de ICS-por partes (ICSPP) [2], que también se denomina método de selecciona-divide-y-conquista para cálculos de IC. A continuación, se desarrollan las ideas principales de ICSPP.

Brevemente, ICSPP consiste en resolver el problema planteado en (4.2) llevando a cabo la división del espacio de configuraciones S en d subespacios, denominados $S_0, S_1, S_2, \dots, S_r$. Los componentes de cada subespacio son seleccionados en base a los mismo criterios de ICS. De una manera tal que el espacio formado por la unión de S_0 y S_1 (que llamaremos espacio S_{01} , cuya dimensión es igual a $d_0 + d + 1$) contenga el mayor número posible de configuraciones que contribuyen de manera más importante a la energía. El resto de los subespacios, que denominaremos S_r , están conformados con configuraciones cuyas contribuciones energéticas sean cada vez menores.

Además, el tamaño de cada uno de los subespacios se elige de acuerdo a criterios relacionados tanto con la cantidad de los recursos de hardware que se dispone como en una estimación del tiempo que tomará llevar a cabo el cálculo. De tal forma que dicha selección permita obtener el mejor resultado posible en un período de tiempo adecuado.

Una vez seleccionado el subespacio S_{01} , se procede a resolver el problema de valores propios dentro de éste subespacio:

$$H^{S_0+S_1} C^{0+1} = E^{0+1} C^{0+1} \quad (4.3)$$

Después de terminado el cálculo, el vector de coeficientes en S_1 (la parte que corresponde a los valores de la solución de (4.2) en el subespacio S_1) se congela, es decir, el valor de dichos coeficientes se mantiene fijo durante el resto del cálculo, lo que es equivalente a contraer dicho subespacio a un espacio unidimensional, u_1 .

Posteriormente, se procede a agregar los términos que corresponden al subespacio S_2 , de manera que tenemos un espacio de dimensión $d_0 + 1 + d_2$, cuyo problema de valores propios está dado por:

$$H^{S_0+u_1+S_1} C^{0+u_1+2} = E^{0+1} C^{0+1} \quad (4.4)$$

El problema de valores propios (4.4) se resuelve y se repite lo hecho en el primer paso, esto es, se congelan los coeficientes de S_2 y se agregan los términos correspondientes a S_3 . El proceso se repite hasta incluir a todos los subespacios de S .

4.1. Criterios de selección de los subespacios de S .

4.1.1. Criterios de Energía y pseudoenergía.

Mediante el uso de la fórmula de Brown, ecuación (3.9), se lleva a cabo el proceso de selección de las configuraciones de simetría normalizadas, G_k , desconectadas que formarán parte de cada uno de los subespacios de S .

De manera tal que el subespacio S_a ($a = 0, 1, 2, \dots, R$) debe incluir todas aquellas G_k desconectadas, cuya contribución a la energía sea mayor que el valor de cierto umbral de energía, $T_a^{\text{energía}}$:

$$|\Delta E_K| > T_a^{\text{energía}}. \quad (4.5)$$

Además de imponer la condición de que los umbrales sean cada vez menores:

$$T_0^{\text{energía}} \geq T_1^{\text{energía}} \geq T_2^{\text{energía}} \geq \dots \geq T_R^{\text{energía}} \quad (4.6)$$

Por otra parte las configuraciones desconectadas de cada S_a son escogidas en base al criterio de pseudoenergía que representa la ecuación (3.26). De tal manera que el criterio de selección viene dado por:

$$|\Delta E_K^{\text{pseudoenergía}}| > T_a^{\text{pseudoenergía}}. \quad (4.7)$$

con la condición adicional:

$$T_0^{\text{pseudoenergía}} \geq T_1^{\text{pseudoenergía}} \geq T_2^{\text{pseudoenergía}} \geq \dots \geq T_R^{\text{pseudoenergía}} \quad (4.8)$$

Capítulo 5

Estado fundamental del H₂O.

5.1. Antecedentes

El método de IC es la manera más directa y general para el tratamiento de la correlación electrónica de átomos y moléculas[6]. En particular, el ICC representa la solución exacta a la ecuación de Schrödinger dentro de una base dada y es de interés general debido a que mediante ella se pueden llevar a cabo cálculos de valores de referencia cuya finalidad sea la calibración de otros métodos más aproximados [5]. Sin embargo, dichos esfuerzos solo se han llevado a cabo utilizando bases orbitales demasiado pequeñas, debido a que el número total de configuraciones crece de manera exponencial.

Los métodos de ICS y de ICSP han demostrado ser alternativas útiles para aproximar energías de IC utilizando hasta excitaciones decuples en sistemas atómicos [1, 2, 44]. Una de las características de ICS es la posibilidad para obtener soluciones numéricas de la ecuación de Schrödinger con una exactitud predeterminada, lo cual es uno de los objetivos de una teoría de estructura electrónica[4].

En este trabajo, se reporta la primera implementación del ICS y ICSP para sistemas moleculares. Esta nueva funcionalidad se ha incorporado a las últimas versiones de los programas de estructura electrónica: ATMOL y AUTOCL. En este sentido, es necesario mencionar que las integrales moleculares calculadas sobre las funciones primitivas se obtuvieron a partir de una versión del programa modificada del programa NWChem[47].

Con dichas integrales el programa ATMOL realiza la transformación de cuatro índices para obtener las integrales electrónicas sobre orbitales de moleculares de simetría y con las cuales es posible llevar a cabo los correspondientes cálculos de ICS.

Se eligió el estado fundamental del H₂O porque este sistema se ha convertido en sistema de referencia dentro de los métodos de mecánica de primeros principios[16]. Aunque lo óptimo hubiera sido realizar todos los cálculos con los conjuntos de base aug-cc-pCVnZ, lamentablemente no hemos sido capaces de llevar a cabo los cálculos con una base aug-cc-pCV6Z. Ya que para llevar a cabo la transformación de 4-índices de sus 533 funciones primitivas las versiones actuales de nuestro programa necesitarían tener acceso a más de 350 GB de RAM. Por lo cual, este estudio se basa principalmente en los resultados obtenidos con las bases cc-pCVnZ. El conjunto de orbitales más grande que utilizamos fue cc-pCV6Z, el cual incluye 412 funciones primitivas.

Los errores de truncamiento van desde menos de 1 microHartree para las bases doble-zeta hasta menos 100 microHartree para la base cc-pCV6Z. Con base en los resultados obtenidos con bases doble zeta y hasta cuádruple zeta del estado fundamental del H₂O [3, 10, 16], así como en los resultados obtenidos para el Ne (isoelectrónico al H₂O) [2], es de esperarse que la contribución a la energía de las excitaciones mayores a orden seis sean considerablemente menores al 100 microHartrees. De esta manera, se decidió no incluir en este estudio cálculos que contengan excitaciones séptuples, óctuples, nónples y decuples. A pesar de esto, se debe hacer notar que los resultados de ICS hasta séxtuples son tan buenos como los que se obtendrían al realiza cálculos de ICC con un error menor a 160 microHartree (0.1 kcal/mol).

5.2. Resultados y discusión.

Todos los cálculos en este estudio se llevaron a cabo para la molécula del agua en su geometría de equilibrio, $R_{OH} = 95.72$ pm y ángulo H-O-H = 104.52° dentro del grupo de simetría C_{2v} .

Los cálculos CASD, CASD(T) fueron llevados a cabo utilizando el programa NWChem [47]. Los cálculos de ICS que reproducen respectivamente los valores de ICSD, ICSDT, ICSDTQ, ICSDTQQn e ICSDTQQnSx fueron llevados a cabo usando ATMOL y AUTOCL acoplados a una versión modificada de NWChem, del cual se obtienen las funciones de onda SCF y las listas de integrales de uno y dos electrones calculadas sobre el conjunto de funciones primitivas.

Utilizamos los conjuntos de base estándar aug-cc-pCVnZ y cc-pCVnZ tal y como se encuentran en la base de datos de NWChem que son los mismos que pueden ser obtenidos de la página web (<https://bse.pnl.gov/bse/portal>).

Con el fin de validar la implementación, el primer paso fue reproducir los valores de referencia para el estado fundamental de la molécula de agua [3, 16]. Primero, reprodujimos todos los valores de IC reportados por Olsen *et al.*, a la distancia de equilibrio, utilizando el conjunto de orbitales cc-pVDZ (24 orbitales) con resultados que en todos los casos presentaban errores de truncamiento menores a 1 microHartree. Posteriormente, reprodujimos los valores de ICSD, ICSDT, ICSDTQ reportados por Chan y Head-Gordon utilizando una base de 41 orbitales (TZ2P) y además calculamos otras energías de IC que reportamos en la Tabla 5.1.

Chan y Head-Gordon reportan como mejor cota superior a la energía un valor de -76.31471 Hartree y un valor extrapolado del ICC de -76.314715 ± 0.000010 Hartree, nosotros reportamos como mejor cota superior un valor de -76.314711 Hartree con un error de truncamiento de 5 microHartree que concuerda de manera excelente con el valor de Chan y Head-Gordon.

Método	Chan y Head-Gordon [16]	Este trabajo
ICSDTQ	-76.314001	-76.314001
ICSDTQQn		-76.314511(1)
CASDTQ	-76.314696	
ICSDTQQnSx		-76.314701(5)
ICSDTQQnSxSp		-76.314709(5)
DMRG	-76.31471	
ICSDTQQnSxSpOc		-76.314711(5)
ICC	-76.314715(10)	-76.314711(5)

TABLA 5.1: Energías totales (en Hartree) para el estado fundamental de H₂O en una bases RADZ

Nuestra cota superior para la base RADZ fue calculada utilizando un espacio modelo de 18.5×10^6 de CSF obtenidas por selección de un espacio completo de 9×10^9 CSF. Este cálculo tardó menos de 12 horas en un nodo sencillo de 16 procesadores Xeon E5-2670 con 64 GB de RAM y utilizando menos de 160 GB de disco duro para el almacenamiento de la matriz Hamiltoniana.

En la Tabla 5.1 podemos observar que el valor de la energía de ICSDTQQnSx reproduce el valor de ICC con un error menor de 10 microHartree. Este resultado, junto con los valores obtenidos en el caso del átomo de Ne, donde las excitaciones séptuples a décuples contribuyen en menos de 30 microHartree a la energía total del sistemas, nos permiten decir que nuestros valores de ICSDTQQnSx reproducen los valores de ICC con un error menor a 10 microHartree.

Para la base doble zeta los errores reportados son aquellos que resultan directamente de la evaluación de la fórmula de Brown. Para las base más grandes el error reportado es resultado de un análisis de sensibilidad. Para el conjunto de bases aug-cc-pCVnZ y cc-pVnZ la energía de SCF extrapolada a base infinita puede ser obtenida mediante la fórmula [10, 48, 49]:

$$E_{SCF}(X) = E_{SCF}(CBC) + ae^{-bX} \quad (5.1)$$

Mediante el tratamiento estadístico de los resultados mostrados en la Tabla 5.2 pudimos evaluar los valores de SCF a base infinita, que se encuentran reportados en la misma tabla.

Conjunto base	SCF/RHF	Conjunto base	SCF/RHF
cc-pCVDZ	-76.027204	aug-cc-pCVDZ	-76.041725
cc-pCVTZ	-76.057358	aug-cc-pCVTZ	-76.060839
cc-pCVQZ	-76.064948	aug-cc-pCVQZ	-76.066094
cc-pCV5Z	-76.067105	aug-cc-pCV5Z	-76.067330
cc-pCV6Z	-76.067411	aug-cc-pCV6Z	-76.067466
Ajustando a la Ec. (5.1)			
Límite estimado (3-6)	-76.0677		-76.0676
Límite estimado (4-6)	-76.0675		-76.0675

TABLA 5.2: Energías SCF para el H₂O en estado fundamental.

Para la energía de correlación se ha encontrado que la convergencia a base infinita es mucho más lenta que en el caso de la energía de SCF, formalmente, cercana al inverso de la tercera potencia [10, 48, 49]:

$$E_{corr}(X) = E_{corr}(CBC) + aX^{-3}, \quad (5.2)$$

Otra fórmula muy utilizadas en cálculos atómicos es:

$$E_{corr}(X) = E_{corr}(CBC) + a(X + b)^{-3}, \quad (5.3)$$

la cual, a su vez es un caso especial de la fórmula general:

$$E_{corr}(X) = E_{corr}(CBC) + a(X + b)^{-3+c}, \quad (5.4)$$

Utilizada por Helgaker *et al.* [48] Sin embargo, en nuestro estudio el mejor ajuste a los datos obtenidos se logró mediante el uso de la siguiente fórmula [50, 51]:

$$E_{corr}(X) = E_{corr}(CBC) + aX^{-3} + bX^{-4}, \quad (5.5)$$

Las Tablas 5.3 y 5.4 contienen nuestros valores calculados de las energías de correlación para cada una de las bases, utilizando ya sea cúmulos acoplados o interacción de configuraciones. Se observa que conforme aumenta el tamaño de la base lo valores de CASD(T) se van desplazando por debajo de nuestros valor de ICSDTQ, lo cual podría indicar que CASD(T) sobre estima la energía de correlación.

Tambien podemos observar que el error en la extrapolación aumenta gradualmente de 0.6 en el caso de ICSD a 0.9 miliHartree para el caso de ICSDTQ y esta diferencia se mantiene constante hasta ICSDTQQnSx.

Conjunto base (n)	ICSD	ICSdT	ICSdTQ	ICSdTQ _{Qn}	ICSdTQ _{QnSx}	CASD	CASD(T)
cc-pCVDZ	-0.241765	-0.244849	-0.255271	-0.255484	-0.255661	-0.251750	-0.255038
cc-pCVTZ	-0.310648	-0.317919	-0.33203	-0.3326	-0.3328	-0.324186	-0.332421
cc-pCVQZ	-0.331820	-0.340055	-0.3554	-0.3560	-0.3562	-0.346496	-0.356018
cc-pCV5Z	-0.338901	-0.347465	-0.3631	-0.3637	-0.3640	-0.353947	-0.363944
cc-pCV6Z	-0.341671	-0.35034	-0.3660	-0.3667	-0.3669	-0.356832	-0.366990
Ajuste a la Ec. (5.5)							
Límite estimado(3-6)	-0.3454	-0.3542	-0.3700	-0.3706	-0.3709	-0.3607	-0.3712
Límite estimado(4-6)	-0.3448	-0.3535	-0.3691	-0.3697	-0.3700	-0.3600	-0.3704

TABLA 5.3: Energías de correlación para el H₂O en estado fundamental (Hartree) y los valores extrapolados, para la familia de bases orbitales cc-pCVnZ

Varios estudios señalan que para obtener mejores valores extrapolados es necesario utilizar valores de n mayores a 3 y que cada vez que sea posible lo mejor es incluir una base con n = 6 o mayor, de esta manera para llevar a cabo el tratamiento estadístico eliminamos los valores de todas los conjuntos de base doble zeta y no llevamos a cabo las extrapolaciones con las base aumentadas aug-cc-pCVnZ debido a que no hemos podido obtener los valores calculados, al no poder llevar a cabo la transformación de cuatro índices de sus 420 funciones primitivas.

Conjunto base (n)	ICSD	ICSdT	ICSdTQ	ICSdTQ _{Qn}	ICSdTQ _{QnSx}	CASD	CASD(T)
aug-cc-pCVDZ	-0.256091	-0.261058	-0.273341	-0.27374	-0.27397	-0.267812	-0.273298
aug-cc-pCVTZ	-0.315388	-0.32341	-0.3383	-0.3389	-0.3391	-0.329593	-0.338797
aug-cc-pCVQZ	-0.333710	-0.34220	-0.3578	-0.3584	-0.3587	-0.348645	-0.358541
aug-cc-pCV5Z	-0.339688	-0.34832	-0.3640	-0.3646	-0.3649	-0.354810	-0.364936

TABLA 5.4: Energías de correlación para el H₂O en estado fundamental (Hartree) y los valores extrapolados, para la familia de bases orbitales aug-cc-pCVnZ

De esta manera nuestra mejor estimación al valor de ICC con la familia de bases cc-pCVnZ se obtiene de la siguiente manera:

$$E(ICC) = -76.0677 (SCF - CBC) - 0.3709 (ICSdTQ_{QnSx} - CBC) = -76.4386 \text{ (9) Hartree.}$$

El cual presenta un error de casi 1 miliHartree, el cual es muy superior al error absoluto que tiene cada uno de nuestras cotas superiores a la energía y el cual proviene esencialmente del error que encontramos al llevar a cabo la extrapolación a base infinita.

Con el fin de disminuir este error llevamos a cabo distintos análisis de los datos, dentro de los cuales el que mejor resultado nos dio fue el análisis de la contribución a la energía de correlación por nivel de excitación para cada uno de los conjuntos de base, esto es, la contribución a la energía de las excitaciones dobles es la diferencia de energía entre el ICSD y el SCF, para las excitaciones triples es la diferencia de energía entre el ICSdT y el ICSDT y así sucesivamente, todas estas diferencias de energía se encuentran en la Tabla 5.5.

Conjunto base	ICSD	ICSdT	ICSdTQ	ICSdTQQ _n	ICSdTQQ _n S _x
cc-pCVDZ	-0.241765	-0.00308	-0.01042	-0.00021	-0.00018
cc-pCVTZ	-0.310648	-0.00726	-0.01412	-0.00054	-0.00024
cc-pCVQZ	-0.331820	-0.00824	-0.01532	-0.00060	-0.00026
cc-pCV5Z	-0.338901	-0.00856	-0.01564	-0.00062	-0.00027
cc-pCV6Z	-0.341671	-0.00867	-0.01570	-0.00063	-0.00027
Límite estimado	-0.3454	-0.0088	-0.0157	-0.0006	-0.0003

Tabla 5.5: Contribución a la energía de correlación para el H₂O en estado fundamental (Hartree).

Con lo cual podemos llevar a cabo nuestra segunda estimación a la energía total norelativista de la molécula de agua:

$$\begin{aligned}
 E(ICC) &= -76,0677(SCF - CBC) - 0,3454(SD - CBC) \\
 &\quad -0,0088(ICSDT - extrapolado) - 0,0157(ICSdTQ - convergido) \\
 &\quad -0,0006(ICSdTQQ_n - convergido) - 0,0003(ICSdTQQ_nS_x - convergido) \\
 &= -76,4385(6)Hartree.
 \end{aligned}$$

Con lo cual redujimos el error en nuestra estimación y encontramos una peculiaridad, para encontrar la contribución a la energía de las excitaciones quintuples y sextuples no es necesario utilizar un conjunto de base mayor a una base cuádruple zeta, lo cual se encuentra justifica en el análisis de sensibilidad que llevamos a cabo en el SCI-TE.

El error reportado es aquel que esta asociado con la extrapolación a base infinita del IC. Y este resultado concuerda de muy buena manera con las extrapolaciones a base infinita “puras” y también reducen la incertidumbre asociada con la inclusión de las excitaciones cuádruples a sextuples al llevar a cabo la extrapolación

5.3. Conclusiones.

Logramos obtener la mejor cota superior reportada en la literatura a la energía del estado fundamental de la molécula de agua, -76.4343 Hartree, a nivel de ICSdTQQ_nS_x, con una base cc-pCV6Z. Este valor recupera más del 98.8% de la energía de correlación del sistema y se encuentra a menos de 3 kcal/mol arriba del valor “experimental”.

Los valores de ICSDTQQnSx calculados con cada una de las bases presenta un error estimado *a priori* menor a 0.1 kcal/mol con respecto al correspondiente ICC. A pesar de esto, los errores asociados a las extrapolaciones a CBC son considerablemente más grandes, lo cual no permite estimar el valor de ICC a CBC, con la exactitud deseada.

Con los valores obtenidos llevamos a cabo extrapolaciones a CBC, encontrando un valor estimado de la energía no relativista del agua de -76.4386(9) Hartree. Analizando los patrones de convergencia de la contribución a la energía por orden de excitación obtuvimos un valor estimado de -76.4386(6) Hartree.

Con una base cuádruple zeta recuperamos de manera cuantitativa la contribución a la energía debida a las excitaciones quintuples y séxtuples.

Para obtener mejores extrapolaciones al ICC-CBC se necesita estimar de mejor manera el ICSD y ICSDT, ya que la contribución a la energía de las excitaciones cuádruples en adelante ya han convergido (a nivel de 0.1 miliHartree).

Capítulo 6

Disociación simétrica del H₂O.

6.1. Antecedentes

Uno de los problemas abiertos de más importantes en la química cuántica es el desarrollo de métodos capaces de calcular superficies de energía potencial (SEP) confiables [5]. La principal dificultad que se presenta en este tipo de estudios es la aparición de diferentes tipos de (cuasi)-degeneraciones en diferentes lugares de las SEP [52–54].

El ejemplo más común ocurre cuando una sola configuración es suficiente para describir de manera adecuada a una molécula cerca de su geometría de equilibrio, pero cuando se trata de representar la formación o ruptura de un enlace otra u otras configuraciones pueden ganar peso dentro de la función de onda. Estas cuasi-degeneraciones causan que la energía de correlación cambie de manera radical en diferentes puntos de una SEP. Un buen método químico cuántico debería ser capaz de recuperar la energía de correlación de con la misma exactitud en todos los puntos de una SEP [54].

Esta situación hace necesario la utilización de métodos químico cuánticos que incorporen excitaciones de orden superior (triples, cuádruples, etc.) en la función de onda. En teoría esta sola condición sería necesaria para obtener energías cada vez más exactas, sin embargo, la experiencia muestra que únicamente los métodos variacionales son capaces de producir resultados sistemáticamente más exactos [3]. La manera más práctica de incluir las excitaciones de orden superior es mediante el uso de los métodos multirreferenciales.

Un gran número de resultados multirreferenciales han sido reportados para el proceso de disociación simétrica de la molécula de agua [55–57]. Aunque éste proceso es totalmente artificial, se eligió estudiarlo porque representa un enorme reto simular la ruptura simétrica de ambos enlaces O-H y se ha convertido en uno de los sistemas modelos para evaluar el desempeño de cualquier método multirreferencial.

Sin embargo, la aproximación más realista sería estudiar la disociación homolítica y heterolítica de uno solo de los enlaces H-O, para los cuales existen una gran variedad de datos tanto teóricos como experimentales [58].

Con el fin de llevar a cabo la tarea de la manera más práctica se preparó un conjunto de archivos de entrada generales que permitirían utilizar nuestra metodología de manera sistemática sin la necesidad de la manipulación por parte del usuario.

6.2. Resultados y discusión

Para comenzar reproducimos todos los valores de IC truncado y de ICC reportados por Olsen *et al.* [3]. Esta no fue tarea sencilla porque los valores de HF reportados por dichos autores no son mínimos globales del SCF-HF, por lo cual es necesario llevar a cabo una manipulación cuidadosa de los orbitales de HF y los parámetros del eigenresolutor de NWChem y PSI4 con el fin de obtener dichas funciones de HF.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
SCF	-76.024039	-75.802387	-75.587711	-75.441244	-75.344392	-75.281577	-75.241515	-75.177814	-75.154299
ICSD	-76.229837	-76.043630	-75.879650	-75.789996	-75.746998	-75.731194	-75.726424	-75.722545	-75.721491
ICSDT	-76.232817	-76.050891	-75.895571	-75.817309	-75.793827	-75.788250	-75.786193	-75.784069	-75.783533
ICSDTQ	-76.241533	-76.070957	-75.945848	-75.905645	-75.895802	-75.892683	-75.891516	-75.890638	-75.890530
IC-5	-76.241721	-76.071806	-75.949431	-75.913160	-75.905514	-75.903406	-75.902654	-75.902139	-75.902092
IC-6	-76.241857	-76.072333	-75.951608	-75.917870	-75.911787	-75.910537	-75.910225	-75.910112	-75.910111
IC-7	-76.241859	-76.072345	-75.951658	-75.917972	-75.911921	-75.910686	-75.910380	-75.910270	-75.910269
IC-8	-76.241859	-76.072347	-75.951667	-75.917991	-75.911945	-75.910713	-75.910409	-75.910301	-75.910299
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410	-75.910302	-75.910301
CASD	-76.238116	-76.062305	-75.929633	-75.897684	-75.901097	-75.905310	-75.907687		
CASD(T)	-76.241202	-76.070717	-75.955486	-75.960554	-76.002459	-76.029397	-76.043487		
CASDT	-76.241367	-76.070925	-75.953070	-75.942743	-75.952072	-75.957345	-75.959914		
CASDT(2) _Q	-76.241792	-76.072066	-75.952195	-75.925579	-75.920325	-75.913081	-75.904487		
CASDTQ	-76.241841	-76.072227	-75.951635	-75.920352	-75.916679	-75.916499	-75.916622		
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410	-75.910302	-75.910301

TABLA 6.1: Energías totales en Hartree para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados por Olsen *et al.* [3]

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
ICSD	0.012023	0.028718	0.072017	0.127995	0.164948	0.179520	0.183986	0.187756	0.188809
ICSDT	0.009043	0.021457	0.056096	0.100682	0.118119	0.122465	0.124218	0.126233	0.126767
ICSDTQ	0.000327	0.001390	0.005819	0.012346	0.016144	0.018032	0.018894	0.019663	0.019771
IC-5	0.000139	0.000542	0.002236	0.004831	0.006432	0.007309	0.007756	0.008163	0.008209
IC-6	0.000003	0.000014	0.000059	0.000121	0.000159	0.000177	0.000185	0.000190	0.000190
IC-7	0.000001	0.000003	0.000009	0.000019	0.000025	0.000029	0.000030	0.000031	0.000032
IC-8	0.000001	0.000000	0.000000	0.000000	0.000001	0.000001	0.000001	0.000001	0.000001
CASD	0.003744	0.010043	0.022034	0.020307	0.010849	0.005404	0.002723		
CASD(T)	0.000658	0.001630	-0.003819	-0.042563	-0.090513	-0.118683	-0.133076		
CASDT	0.000493	0.001423	-0.001403	-0.024752	-0.040126	-0.046631	-0.049504		
CASDT(2) _Q	0.000068	0.000282	-0.000528	-0.007588	-0.008379	-0.002367	0.005923		
CASDTQ	0.000019	0.000121	0.000032	-0.002361	-0.004733	-0.005784	-0.006211		

TABLA 6.2: Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.1

En la Tabla 6.1 se muestran los valores de energía obtenidos mediante ICS, todos estos valores coinciden con los valores reportados por Olsen *et al.* con un error máximo de 5 microHartrees. Dichos autores solo reportan valores de energía para el sistema con una distancia O-H igual a 3 veces R_{eq} . Nosotros decidimos ir más allá y calculamos los valores de energía de IC y CA para distancias hasta 4 veces R_{eq} utilizando NWChem y PSI4. Sin embargo, no fue posible llevar a cabo cálculos a distancias internucleares mayores debido a que no pudimos encontrar soluciones de RHF a 6 R_{eq} y 8 R_{eq} con PSI4.

Con ICS y NWChem pudimos realizar cálculos hasta 8 R_{eq} . Al calcular estos valores de IC y CA seguimos el mismo proceso utilizado para obtener los valores reportados por Olsen *et al.* Lo más importante del proceso de obtención de la función de onda SCF-HF es preservar la constitución del vector de IC, es decir, que las funciones de onda obtenidas presentan las mismas configuraciones dominantes durante el proceso de disociación.

En la Tabla 6.2 se muestran diferencias de energía con respecto a las energías de ICC para cada uno de los métodos presentados en la Tabla 6.1. Al analizar la Tabla 6.2 podemos observar que conforme aumenta la distancia H-O los métodos de cúmulos acoplado tienden sistemáticamente a obtener valores de energía por debajo de los correspondientes energías de ICC. Olsen *et al.* justifican este fenómeno por el hecho de que estos resultados están basados en funciones de onda SCF-RHF (que no es consistente en tamaño). Lo anterior lo “remedian” ligeramente mediante la utilización de funciones de onda UHF (consistente en tamaño).

Sin embargo, la principal razón de éstas fallas parece radicar en el hecho de que cuando ocurren cuasi-degeneraciones en la función de onda, las amplitudes de las configuraciones desconectadas de CA pueden resultar demasiado grandes [59, 60], lo cual se puede corregir al incluir un conjunto adecuado de configuraciones conectadas de mayor orden de excitación en la función de onda de CC [61, 62].

Es importante señalar que a pesar de que los resultados de CA no sean cotas superiores a la energía del sistema Olsen *et al.* remarcan en su artículo la “superioridad” de CA sobre IC al recuperar de manera más “eficiente” la energía de correlación para un mismo orden de excitación. De esta manera señalan que a 2 R_{eq} CASDT recupera más energía de correlación que ICSDT e ICSDTQ, sin notar que ese mismo valor se encuentra por debajo del resultado de ICC. Por otra parte, cuando utilizamos una técnica IC el hecho de utilizar funciones SCF-RHF no representa ningún problema y se obtienen soluciones que mejoran de manera sistemática.

R_{OH}	1	1.5	2	2.5	3	3.5	4
SCF	-76.024039	-75.802387	-75.587711	-75.469758	-75.434945	-75.421500	-75.414283
ICSD	-76.229837	-76.043630	-75.879650	-75.779122	-75.769271	-75.775299	-75.777491
ICSDT	-76.232817	-76.050891	-75.895571	-75.860331	-75.845749	-75.841282	-75.840233
ICSDTQ	-76.241533	-76.070957	-75.945848	-75.910378	-75.906945	-75.906994	-75.907074
IC-5	-76.241721	-76.071806	-75.949431	-75.916999	-75.911331	-75.910130	-75.909831
IC-6	-76.241857	-76.072333	-75.951608	-75.917938	-75.911916	-75.910695	-75.910394
IC-7	-76.241859	-76.072345	-75.951658				
IC-8	-76.241859	-76.072347	-75.951667				
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410

TABLA 6.3: Energías totales de IC directo (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos usando PSI4.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
SCF	-76.024039	-75.802387	-75.587711	-75.469758	-75.434945	-75.421500	-75.414283	-75.400153	-75.393278
ICSD	-76.229837	-76.043630	-75.879649	-75.779122	-75.769271	-75.775299	-75.777491	-75.777896	-75.777836
ICSDT	-76.232817	-76.050891	-75.895571	-75.860331	-75.845749	-75.841282	-75.840233	-75.839968	-75.839966
ICSDTQ	-76.241533	-76.070957	-75.945848	-75.910378	-75.906945	-75.906995	-75.907074	-75.907055	-75.906937
IC-5	-76.241721	-76.071806	-75.949431	-75.916999	-75.911331	-75.910130	-75.909831	-75.909724	-75.909492
IC-6	-76.241857	-76.072333	-75.951608	-75.917938	-75.911916	-75.910695	-75.910393	-75.910003	-75.909765
IC-7	-76.241859	-76.072345	-75.951658	-75.917988	-75.911944	-75.910712	-75.910406	-75.910007	-75.909769
IC-8	-76.241859	-76.072347	-75.951667	-75.917990	-75.911944	-75.910712	-75.910407	-75.910007	-75.909769
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410	-75.910302	-75.910301

TABLA 6.4: Energías totales de ICS (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSD

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
SCF	-76.024039	-75.802387	-75.587711	-75.469758	-75.434945	-75.421500	-75.414283	-75.400153	-75.393278
ICSD	-76.229837	-76.043630	-75.879649	-75.779122	-75.769271	-75.775299	-75.777491	-75.773819	-75.777874
ICSdT	-76.232817	-76.050891	-75.895571	-75.860331	-75.845749	-75.841282	-75.840233	-75.865398	-75.865569
ICSdTQ	-76.241533	-76.070957	-75.945848	-75.910378	-75.906945	-75.906995	-75.907074	-75.907194	-75.907108
IC-5	-76.241721	-76.071806	-75.949431	-75.916999	-75.911331	-75.910130	-75.909831	-75.910131	-75.910133
IC-6	-76.241857	-76.072333	-75.951608	-75.917938	-75.911916	-75.910695	-75.910394	-75.910302	-75.910300
IC-7	-76.241859	-76.072345	-75.951658	-75.917988	-75.911944	-75.910713	-75.910408	-75.910302	-75.910301
IC-8	-76.241859	-76.072347	-75.951667	-75.917990	-75.911945	-75.910713	-75.910408	-75.910302	-75.910301
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410	-75.910302	-75.910301

TABLA 6.5: Energías totales de ICS (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSdT

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
SCF	-76.024039	-75.802387	-75.587711	-75.469758	-75.434945	-75.421500	-75.414283	-75.400153	-75.393278
ICSD	-76.229837	-76.043630	-75.879649	-75.779122	-75.769271	-75.778813	-75.792869	-75.793776	-75.793776
ICSdT	-76.232817	-76.050891	-75.895571	-75.860331	-75.845749	-75.859558	-75.860098	-75.859954	-75.859953
ICSdTQ	-76.241533	-76.070957	-75.945848	-75.910378	-75.906945	-75.908116	-75.908947	-75.908925	-75.908924
IC-5	-76.241721	-76.071806	-75.949431	-75.916999	-75.911331	-75.910466	-75.910188	-75.910131	-75.910133
IC-6	-76.241857	-76.072333	-75.951608	-75.917938	-75.911916	-75.910708	-75.910409	-75.910302	-75.910300
IC-7	-76.241859	-76.072345	-75.951658	-75.917988	-75.911945	-75.910714	-75.910410	-75.910302	-75.910301
IC-8	-76.241859	-76.072347	-75.951667	-75.917991	-75.911945	-75.910714	-75.910410	-75.910302	-75.910301
ICC	-76.241860	-76.072348	-75.951667	-75.917991	-75.911946	-75.910714	-75.910410	-75.910302	-75.910301

TABLA 6.6: Energías totales de ICS (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSdTQ

R_{OH}	1	1.5	2	2.5	3	3.5	4
ICSD	0.012023	0.028718	0.072017	0.138869	0.142675	0.135416	0.132920
ICSdT	0.009043	0.021457	0.056096	0.057660	0.066197	0.069432	0.070177
ICSdTQ	0.000327	0.001390	0.005819	0.007613	0.005001	0.003720	0.003337
IC-5	0.000139	0.000542	0.002236	0.000991	0.000615	0.000584	0.000580
IC-6	0.000003	0.000014	0.000059	0.000053	0.000030	0.000019	0.000016
IC-7	0.000001	0.000003	0.000009				
IC-8	0.000001	0.000000	0.000000				

TABLA 6.7: Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.3.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
ICSD	0.012023	0.028718	0.072018	0.138869	0.142675	0.135415	0.132920	0.132406	0.132465
ICSdT	0.009043	0.021457	0.056096	0.057660	0.066197	0.069432	0.070177	0.070334	0.070335
ICSdTQ	0.000327	0.001390	0.005819	0.007613	0.005001	0.003720	0.003337	0.003247	0.003364
IC-5	0.000139	0.000542	0.002236	0.000991	0.000615	0.000584	0.000579	0.000578	0.000809
IC-6	0.000003	0.000014	0.000059	0.000053	0.000030	0.000019	0.000017	0.000299	0.000535
IC-7	0.000001	0.000003	0.000009	0.000003	0.000002	0.000003	0.000004	0.000295	0.000532
IC-8	0.000001	0.000000	0.000000	0.000000	0.000002	0.000002	0.000004	0.000295	0.000532

TABLA 6.8: Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.4.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
CISD	0.012023	0.028718	0.072018	0.138869	0.142675	0.135415	0.132919	0.136482	0.132427
CISDT	0.009043	0.021457	0.056096	0.057660	0.066197	0.069432	0.070177	0.044904	0.044731
CISDTQ	0.000327	0.001390	0.005819	0.007613	0.005001	0.003720	0.003337	0.003108	0.003193
IC-5	0.000139	0.000542	0.002236	0.000991	0.000615	0.000584	0.000579	0.000170	0.000168
IC-6	0.000003	0.000014	0.000059	0.000053	0.000030	0.000019	0.000017	0.000000	0.000000
IC-7	0.000001	0.000003	0.000009	0.000003	0.000001	0.000002	0.000002	0.000000	0.000000
IC-8	0.000001	0.000000	0.000000	0.000000	0.000001	0.000002	0.000002	0.000000	0.000000

TABLA 6.9: Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.5.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
CISD	0.012023	0.028718	0.072018	0.138869	0.142675	0.131902	0.117542	0.116526	0.116525
CISDT	0.009043	0.021457	0.056096	0.057660	0.066197	0.051156	0.050313	0.050348	0.050347
CISDTQ	0.000327	0.001390	0.005819	0.007613	0.005001	0.002599	0.001464	0.001376	0.001377
IC-5	0.000139	0.000542	0.002236	0.000991	0.000615	0.000249	0.000222	0.000170	0.000168
IC-6	0.000003	0.000014	0.000059	0.000053	0.000030	0.000006	0.000001	0.000000	0.000000
IC-7	0.000001	0.000003	0.000009	0.000003	0.000001	0.000000	0.000001	0.000000	0.000000
IC-8	0.000001	0.000000	0.000000	0.000000	0.000001	0.000000	0.000001	0.000000	0.000000

TABLA 6.10: Diferencias de energía con respecto a las energías de ICC para el estado fundamental de H₂O con una base cc-pVDZ a distintas distancia internucleares H-O, con los valores de HF reportados en la Tabla 6.6.

Al realizar el presente trabajo se observó que éste tipo de estudios resulta en una considerable carga de trabajo porque para cada uno de los puntos evaluados es necesario llevar a cabo un proceso “casi quirúrgico” para obtener los valores de la función de onda de manera adecuada. Por lo cual se decidió probar que pasa si simplemente se toma la función de onda de SCF-RHF que producen de manera automática NWChem o PSI4.

Ya que PSI4 permite llevar a cabo este tipo de cálculos de manera totalmente automática, se procedió directamente a realizar los cálculos de IC truncados (hasta ICSDTQQnSx) y de ICC con el fin de comparar dichos resultados con los obtenidos por Olsen *et al.* Las cotas superiores a la energía de dichos cálculos se muestran en la Tabla 6.3. Éstos resultados muestran que que ambos procedimientos convergen sistemáticamente a los mismos valores de ICC mostrados en la Tabla 6.1.

Ambos conjuntos de resultados a nivel de ICC presentan los mismos valores propios (del estado fundamental y los estados excitados), así como los mismos números de ocupación de los respectivos orbitales naturales, por lo cual lo más probable es que ambos conjuntos estén relacionados por una transformación ortogonal de sus orbitales.

En la Tabla 6.7 se muestran diferencias de energía con respecto a las energías de ICC para los resultados presentados en la Tabla 6.3. Éstos resultados muestran que los valores de IC obtenidos de ésta manera recuperan una mayor cantidad de energía de correlación que los correspondientes valores obtenidos por Olsen *et al.*, que se muestran en la Tabla 6.2.

Posteriormente se llevaron a cabo los cálculos de ICS con el fin de reproducir los valores obtenidos con PSI4, los resultados se muestran en la Tabla 6.4. Se observa que ambos métodos producen los mismos resultados hasta 4 R_{eq} . Sin embargo, cuando se alarga la distancia de enlace O-H a 6 R_{eq} y 8 R_{eq} se observa que los valores de ICS no reproducen los valores de IC obtenidos con la metodología de Olsen *et al.* (reportados en la Tabla 6.1).

La Tabla 6.8 muestra que en el caso extremo de 8 R_{eq} , las energías de IC hasta séxtuples quedan más de 0.3 miliHartrees por arriba de los valores de la Tabla 6.1 y aún más los valores de ICS hasta óctuples no recuperan el ICC quedando poco más de 0.5 miliHartrees por encima del valor correcto.

Con el fin de mejorar dichos resultados se procedió a repetir los cálculos utilizando orbitales naturales provenientes de un cálculo de ICSDT, los cuales a 6 R_{eq} y 8 R_{eq} mostraron recuperar el valor de ICC a nivel de 1 microHartree con un cálculo de ICS hasta séxtuples. Ya que el sistema es bastante manejable se procedió a dar el siguiente paso y se utilizaron orbitales naturales producto de un ICSDTQ. Los resultados obtenidos se muestran en la Tablas 6.5, 6.6, 6.9 y 6.10.

Los resultados en la Tabla 6.9 muestra que cuando se utilizan orbitales naturales de un IC que incluye hasta excitaciones triples se obtiene como resultado una convergencia sistemática a la energía de ICC. Sin embargo, es más sorprendente lo que ocurre al incluir excitaciones cuádruples, ya que al hacerlo, se obtiene una convergencia a la energía de ICC aún más rápida a partir de 3.5 R_{eq} , con respecto a los resultados en los cuales se incluyen sólo excitaciones triples. Los resultados de incluir excitaciones cuádruples se muestran en la Tabla 6.10.

La explicación que encuentro más convincente para explicar que es lo que sucede es que alrededor de esta distancia el proceso de ruptura de ambos enlaces O-H comienza a involucrar de manera cuantitativa a las configuraciones que darán origen a las funciones de onda de los átomos aislados (oxígeno en estado 3P y a los dos hidrógenos en estados 2S) en las cuales se involucran 4 electrones. De ésta manera una función de onda de ICSDTQ recupera de manera cuantitativa la información relevante de este proceso en la función de onda, lo que se refleja de manera directa en sus orbitales naturales. Sin embargo, los resultados de la Tabla 6.5 muestran que para recuperar los valores de ICC a 6 y 8 R_{eq} no es absolutamente necesaria la obtención de orbitales naturales a nivel de ICSDTQ.

Los resultados obtenidos en las Tablas 6.5 y 6.6 tienen varias repercusiones, entre las cuales podemos mencionar que el uso de orbitales naturales de un ICAC ayudan construir un mejor espacio S_{01} , lo cual es de gran importancia, porque mi experiencia indica que la mayor parte del error de truncamiento de ICSP se proviene de la buena elección de dicho subespacio.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
Tamaño de S_0	9539	14667	20462	29455	32262	31210	29914	28123	28080
Tamaño de S_{01}	1026978	1216347	1006042	857657	678085	591250	545479	497620	487768
Tamaño de S	6629546	6260147	4799077	3538616	2434575	2041325	1855274	1642207	1584965
$E(S_{01}) - E(ICC)$	0.000011	0.000019	0.000062	0.000124	0.000162	0.000182	0.000191	0.000198	0.000198
$E(IC-6) - E(ICC)$	0.000003	0.000015	0.000059	0.000121	0.000159	0.000177	0.000185	0.000190	0.000190

TABLA 6.11: Tamaños de los subespacios S_0 , S_{01} y los espacios totales en los cuales se calcularon los ICS hasta séxtuples y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H_2O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF reportados por Olsen *et al.*

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
Tamaño de S_0	9525	14801	19920	44782	42102	44168	48834	61735	70358
Tamaño de S_{01}	1028919	1220183	982036	810291	621466	543710	512567	299238	284597
Tamaño de S	6610398	6258198	4707582	2937667	1975929	1637743	1446975	649768	560497
$E(S_{01}) - E(ICC)$	0.000011	0.000019	0.000062	0.000056	0.000037	0.000055	0.000105	0.000011	0.000012
$E(IC-6) - E(ICC)$	0.000003	0.000014	0.000059	0.000053	0.000030	0.000019	0.000017	0.000000	0.000000

TABLA 6.12: Tamaños de los subespacios S_0 , S_{01} y los espacios totales en los cuales se calcularon los ICS hasta séxtuples y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H_2O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDT.

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
Tamaño de S_0	9539	14667	20462	43506	40866	62102	50081	61903	70230
Tamaño de S_{01}	1026978	1216333	1006042	801519	611169	542970	359756	301362	285105
Tamaño de S	6629548	6260142	4799082	2910149	1955012	1538346	959207	652934	561036
E(S_{01}) - E(ICC)	0.000011	0.000019	0.000062	0.000056	0.000037	0.000021	0.000014	0.000007	0.000007
E(IC-6) - E(ICC)	0.000003	0.000014	0.000059	0.000053	0.000030	0.000006	0.000001	0.000000	0.000000

TABLA 6.13: Tamaños de los subespacios S_0 , S_{01} y los espacios totales en los cuales se calcularon los ICS hasta séxtuples y de las diferencias de energías con respecto al ICC (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con los valores de HF obtenidos con NWChem y usando orbitales naturales de ICSDTQ.

Las Tablas 6.11, 6.12 y 6.13 muestran los tamaños de los subespacios S_0 , S_{01} y los espacios totales en los cuales se calcularon los ICS hasta séxtuples para cada distancia O-H de los resultados reportados en las Tablas 6.4, 6.5 y 6.6, respectivamente. En éstas podemos ver que el mayor efecto que se obtiene al incluir utilizar los orbitales naturales de ICSDT o ICSDTQ es la disminución del tamaño del espacio S_{01} . Los resultados de las energía obtenida en estos subespacios muestran la eficiencia de ICS para seleccionar dichos subespacios.

Los resultados de la Tabla 6.11 muestran que dentro del espacio S_{01} se obtiene una cota superior a la energía que difiere en menos de 20 microHartrees con el valor exacto del respectivo ICSDTQ_nSx en todos los casos y los valores de ICS hasta séxtuples difieren con respecto a los valores ICC en menos de 200 microHartrees. La Tabla 6.12 muestra que la máxima diferencia entre la energía dentro de S_{01} y el valor de ICSDTQ_nSx es de menos de 90 microHartree a 4 R_{eq} . La Tabla 6.13 muestra que al usar orbitales naturales de un ICSDTQ estas diferencias vuelven a ser menores a 20 microHartrees. En ambos casos las diferencias entre los valores de ICS hasta sextúples y los de ICC son menores a 60 microHartrees.

Podemos observar que el hecho de utilizar los mínimos globales de RHF para producir orbitales naturales de ICSDT o ICSDTQ conlleva a una mejor selección del espacio modelo utilizados para calcular el ICS. En el caso extremo, a 8 R_{eq} los espacios modelos S_{01} para un IC hasta séxtuples de las Tablas 6.12 y 6.13 contienen menos de 600,000 FSNE, y dentro de ellos se recuperan los valores de ICC a nivel de 1 microHartree, mientras que el espacio modelo M de la Tabla 6.11 es 3 veces mayor y queda 200 microHartrees por encima del valor de ICC.

A pesar del excelente rendimiento de este procedimiento hay que recordar que llevar a cabo un cálculo de ICSDTQ se vuelve demasiado costoso rápidamente cuando se considera un sistema con más electrones o se utiliza una base de orbitales más grande, por lo cual ésta manera de optimizar las funciones de onda de referencia no resulta práctica y los resultados muestran que la diferencia con respecto a los orbitales naturales de ICSDT no parecen ser muy grandes.

En caso de ser necesaria la obtención de orbitales naturales a nivel de IC hasta triples o cuádruples una posible alternativa para llevar a cabo este tipo de optimización de la función de onda de referencia de IC es proceder de la siguiente manera:

- 1) obtener la mejor solución SCF posible,
- 2) diagonalizar la matriz densidad a nivel de ICSD,
- 3) obtener un espacio modelo M muy simplificado del ICSDT o ICSDTQ del sistema (incluyendo solo algunas excitaciones triples y cuádruples, manteniendo el tamaño del IC lo más pequeño posible).
- 4) obtener los orbitales naturales de esta nueva función de onda.
- 5) realizar un ICS de manera “habitual”

Otro uso de éste procedimiento sería la identificación de las configuraciones dominantes que se deban incluir al llevar a cabo un desarrollo multirreferencial de la función de onda. Ya que en la mayoría de los casos estas configuraciones aparecen en la función de onda unireferencial como excitaciones triples o cuádruples. De esta manera se tendría un buen método para obtener la función de onda de partida de un cálculo multirreferencial.

La Figura 6.1 muestra los tamaños de los subespacios S_0 , S_1 y los espacios totales en los cuales se calcularon los respectivos valores de ICS hasta séxtuples haciendo uso orbitales naturales de ICSDT como referencia.

Podemos observar que una disminución considerable de los espacio modelo utilizados para representar los correspondientes espacios de IC hasta séxtuples, pasando de más de 7 millones de FSNE cerca de la geometría de equilibrio, a menos de 1 millón de términos cuando el sistema esta completamente disociado, es importante remarcar que los respectivos errores de estos espacios con respecto a los valores de ICC nunca son mayores de 100 microHartrees lo que en éste caso validaría a esta variante como un excelente método para llevar a cabo estudios de disociación molecular al presentar errores con respectos a ICC que varían ligeramente a lo largo de la SEP.

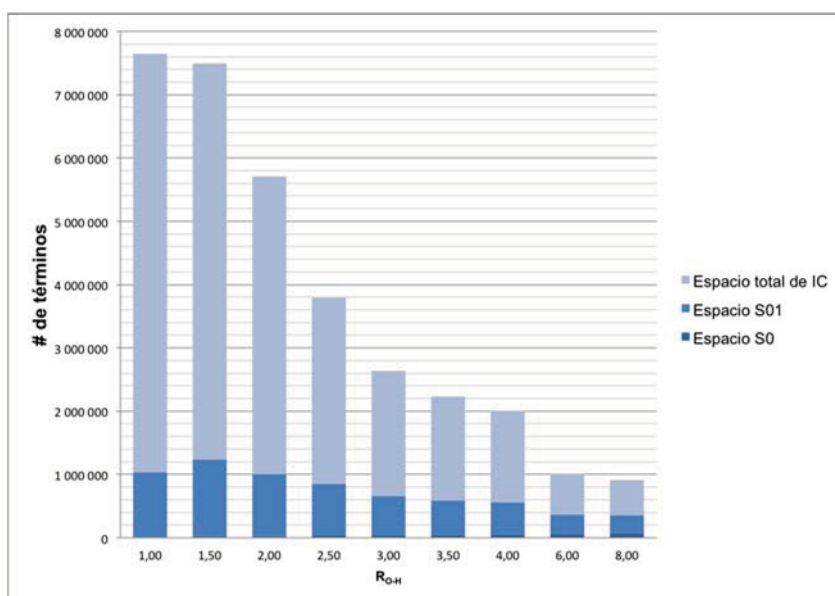


FIGURA 6.1: Tamaños de los subespacios S_0 , S_1 y los espacios totales en los cuales se calcularon los ICS hasta séxtuples con usando orbitales naturales de ICSDT a en función de la distancia H-O.

Al analizar las Tablas 6.1, 6.5 y 6.6 podemos observar que las SEP que se obtienen con los datos de la Tabla 6.1 para cada orden de excitación parecen ser adecuadas para describir de manera correcta el proceso de ruptura de los enlaces. En cada uno de los casos se obtienen SEP lisas y continuas a nivel de las decenas de microHartree, sin embargo, aún en el caso de ICSDTQQnSx los resultados obtenidos continuan presentando un error considerable con respecto a ICC. Por otro, lado en el caso de las PES de las Tablas 6.5 y 6.6 las curvas de ICSD hasta ICSDTQ se vuelven suaves sólo a nivel de 1 miliHartree, aunque, el nivel de exactitud vuelve a restituirse alrededor de 1 microHartree para las curvas de ICSDTQQn.

La exactitud de los cálculos de ICSDTQQnSx de las Tablas 6.5 y 6.6 provoca que nuevamente tengamos curvas bien comportadas y se acercan de tal manera a las curvas de ICC, que están libres de error de consistencia en tamaño a nivel de 1 microHartree.

Con base en éstos resultados, podemos concluir que cualquiera de las dos maneras de abordar el problema planteadas en éste capítulo tienen pros y contras, la aproximación de Olsen *et al.* es la aproximación "tímida" y laboriosa que obtiene resultados de buena calidad que mejoran de manera gradual, y la aproximación de ICS (o mía) es una aproximación que va por todo y que obtiene como resultado una acercamiento mucho más rápido a la energía de ICC, lo que trae como bono extra la disminución del error de consistencia en tamaño.

Algo que hay que remarcar es que en ambos casos los resultados describen de manera adecuada el proceso de ruptura de los enlaces O-H y que sin importar que método utilicemos los valores de la energía tienden de manera sistemática hacia los valores de ICC.

6.3. Comparación con métodos multirreferenciales.

La Figura 6.2 muestra los valores absolutos de las diferencias de energía con respecto al ICC de los diferentes métodos de estructura electrónica reportados en la Tabla 6.2. Podemos observar que éste tipo de aproximación de los métodos de IC muestran un comportamiento predecible de dichos errores.

La Figura 6.3 muestra los valores absolutos de las diferencias de energía con respecto al ICC de los diferentes métodos de estructura electrónica reportados en las Tablas 6.2, 6.9 y 6.10. Podemos observar que las energías obtenidas con orbitales naturales de de ICSDT o ICSDTQ mejoran considerablemente la convergencia a los valores de ICC, disminuyendo con esto los respectivos errores de consistencia en tamaño. Los valores de la energía a ICSDTQ presentan errores mayores a 4 kcal/mol para distancias de enlace entre 2 y 3 veces R_{eq} . Con la inclusión de las excitaciones quintuples los error de los métodos basados en orbitales naturales nunca rebasan los 2 kcal/mol. Al incluir las excitaciones séxtuples, en todos lo casos los errores son menores a 0.15 kcal/mol.

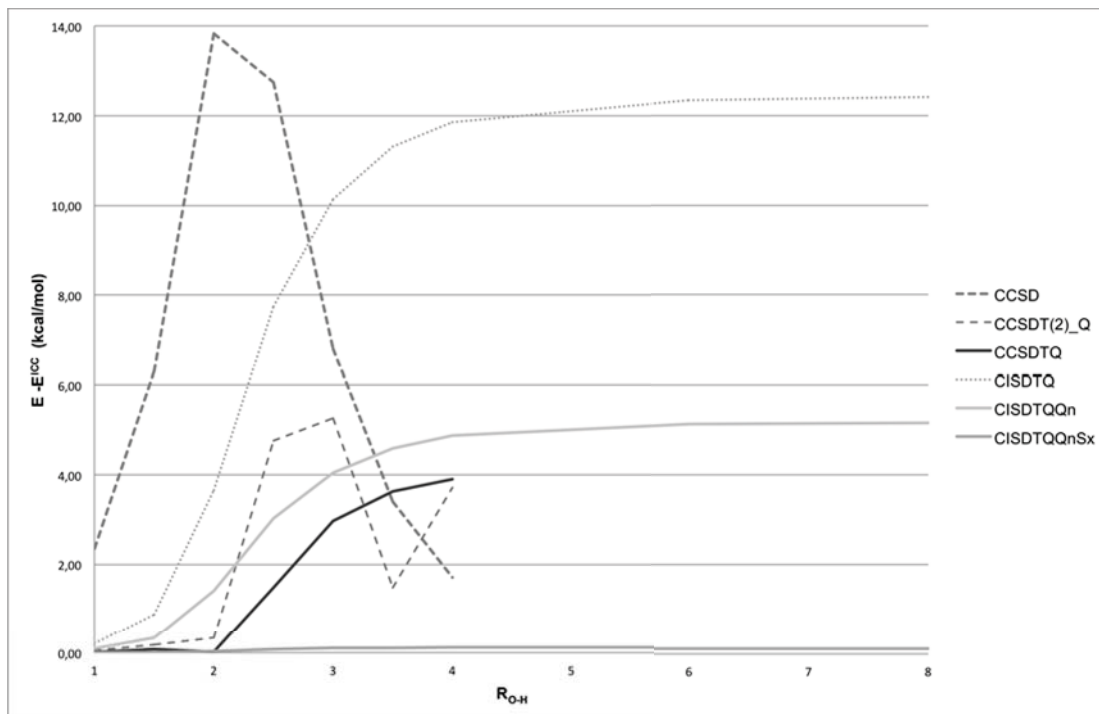


FIGURA 6.2: Valores absolutos de las diferencias de energía con respecto a las energías de ICC en función de la R_{eq} , Para los valores de la Tabla 6.2 (en kcal/mol)

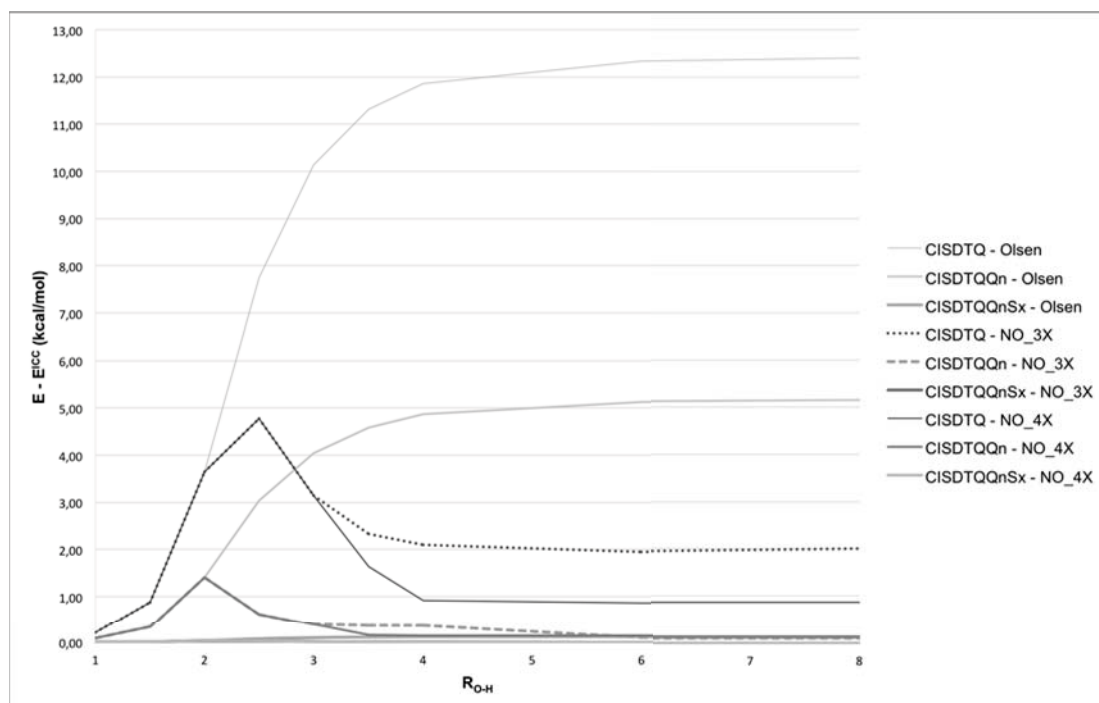


FIGURA 6.3: Valores absolutos de las diferencias de energía con respecto a las energías de ICC en función de la R_{eq} , Para los valores de las Tablas 6.2, 6.9 y 6.10 (en kcal/mol)

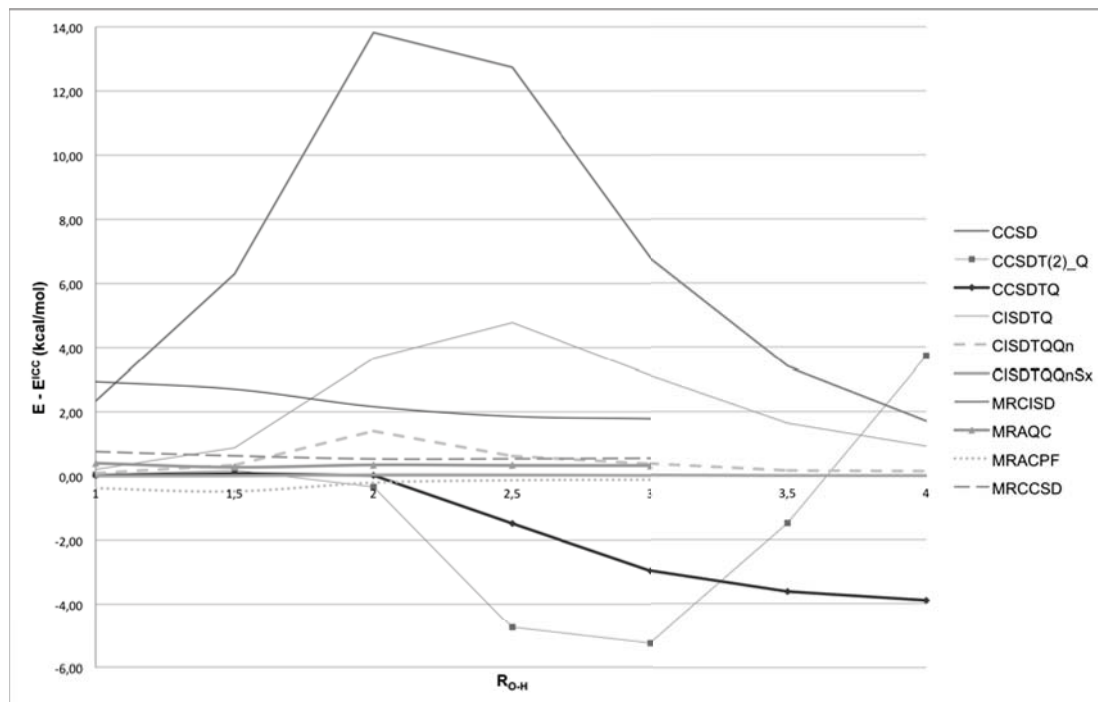


FIGURA 6.4: Valores absolutos de las diferencias de energía con respecto a las energías de ICC en función de la R_{eq} , Para los valores de la Tabla 6.14 (en kcal/mol).

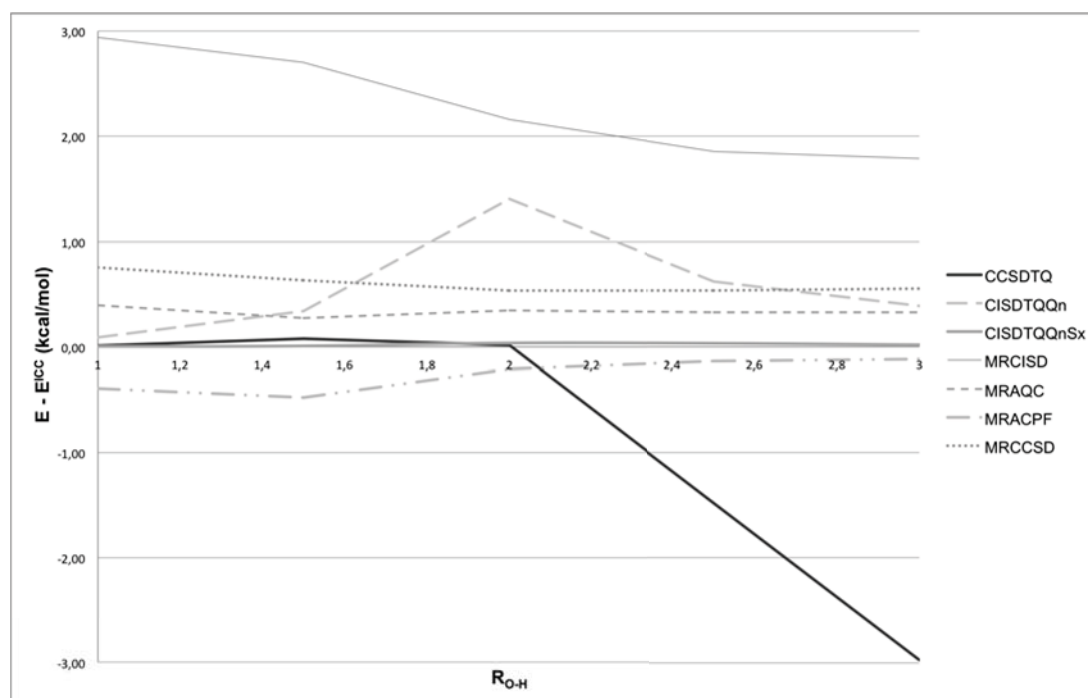


FIGURA 6.5: Valores absolutos de las diferencias de energía con respecto a las energías de ICC en función de la R_{eq} , Para los valores de la Tabla 6.14 (en kcal/mol).

R_{OH}	1	1.5	2	2.5	3	3.5	4	6	8
CASD	0.003744	0.010043	0.022034	0.020307	0.010849	0.005404	0.002723		
CASDT(2) _Q	0.000068	0.000282	-0.000528	-0.007588	-0.008379	-0.002367	0.005923		
CASDTQ	0.000019	0.000121	0.000032	-0.002361	-0.004733	-0.005784	-0.006211		
MRCI	0.004681	0.004309	0.003445	0.002962	0.002847				
MRAQCC	0.000624	0.000436	0.000551	0.000526	0.000520				
MRACPF	-0.000620	-0.000761	-0.000332	-0.000211	-0.000182				
MRCCSD	0.001204	0.001003	0.000849	0.000855	0.000882				
ICSdTQ	0.000327	0.001390	0.005819	0.007613	0.005001	0.002599	0.001464	0.001376	0.001377
IC-5	0.000139	0.000542	0.002236	0.000991	0.000615	0.000249	0.000222	0.000221	0.000221
IC-6	0.000003	0.000014	0.000059	0.000053	0.000030	0.000006	0.000001	0.000001	0.000001
S ₀₁ -Olsen	0.000011	0.000019	0.000062	0.000124	0.000162	0.000182	0.000191	0.000198	0.000197
S ₀₁ -ON-3X	0.000011	0.000019	0.000062	0.000056	0.000037	0.000055	0.000105	0.000011	0.000012

TABLA 6.14: Diferencias de energía con respecto al ICC (en Hartree) para el estado fundamental de H₂O en una base cc-pVDZ a distintas distancia H-O, con distintos métodos multirreferenciales y algunos resultados de las Tablas 6.2, 6.6, 6.11 y 6.12.

La Tabla 6.14 muestra algunos de los valores de las diferencias de energía con respecto a los valores de ICC de las Tablas 6.10, 6.11, 6.12 y valores reportados en la literatura de varios métodos multirreferenciales [55–57]. En las Figuras 6.4 y 6.5 se presentan las diferencias de energía con respecto al ICC de los algunos métodos de estructura electrónica reportados en la Tabla 6.14 en kcal/mol.

En la Figura 6.5 podemos observar que ICS a nivel de excitaciones quintuples supera de manera clara los resultados de MRICSD y compite de manera adecuada con los resultados de MRCCSD. Se observa que ICSdTQ_{QnSx} representa la mejor aproximación a la energía ICC, siendo CASDTQ el único método que esta a su nivel hasta llevar a 2 veces la distancia de equilibrio. Sin embargo, a partir de 2.5 R_{eq} CASDTQ comienza a dar resultados por debajo de los valores de ICC y más allá de 3 R_{eq} no fue posible encontrar soluciones debido a problemas de convergencia.

Se observa que en todos los casos MRACPF reporta valores de energías por debajo de los valores de ICC. MRCCSD, MRAQC y MRACPF presentan excelentes resultados, ya que en todos los casos los errores son menores a 1 kcal/mol.

En la última hilera de la Tabla 6.14 podemos observar la cota superior a la energía calculada dentro del espacio S_{01} (el cual es un subespacio del espacio modelo M, que a su vez es un subespacio del espacio completo de IC, en este caso el espacio de IC hasta séxtuples).

En todos los casos, este valor de energía recupera los valores de ICC con un error menor a 20 microHartrees. En ningún caso el tamaño de este subespacio va más allá de 12000 FSNE. En todos los casos la cota superior a la energía obtenida en este subespacio da mejores resultados que los métodos multirreferenciales más sofisticados.

Este último resultado es importante de remarcar porque una posible variante del método ICS es su uso como modelo de química cuántica. En este trabajo el ICS se ha utilizado para obtener soluciones exactas a la ecuación de Schrödinger, dentro de la exactitud numérica, a nivel molecular esto se considera 1 microHartree. Sin embargo, ICS puede utilizarse para obtener resultado muchos menos exactos para estudiar sistemas más grandes o simplemente para tratar problemas cuyas dependencia de la energía sea considerablemente menor

6.4. Conclusiones.

Se llevó a cabo un estudio de la disociación simétrica de la molécula de agua, partiendo de la geometría de equilibrio hasta 8 R_{eq} . Se siguieron dos metodologías diferentes para la obtención de las funciones de onda de referencia: la primera consiste en seguir un mismo eigenvector de HF a largo de todo el proceso y la segunda en tomar la mejor función de onda de SCF (en el sentido variacional) para cada punto de la superficie de energía potencial.

Los resultados obtenidos con la segunda metodología puso en evidencia la necesidad de utilizar orbitales naturales obtenidos de un IC que incluya excitaciones más allá de simples y dobles. Cuando se usan éste tipo de orbitales naturales se obtiene una convergencia a la energía de ICC mucho más rápida que usando la primera metodología.

Al llevar a cabo la segunda metodología, se desarrolló una manera de optimizar de manera sistemática los orbitales naturales, éste proceso conlleva a la obtención de funciones de onda más compactas que recuperan una mayor cantidad de la energía de correlación. Al llevar a cabo éste proceso los resultados obtenidos presentan errores de consistencia en tamaño considerablemente menores respecto a los resultados de un método más tradicional. La nueva metodología permitiría llevar a cabo una mejor selección de las configuraciones de referencia utilizadas para eventuales cálculos multirreferenciales.

Capítulo 7

Consistencia en tamaño y algunos ejemplos.

7.1. Antecedentes.

Al truncar la función de onda de IC a un determinado orden de excitación, IC deja de ser una teoría exacta [5–9]. Una consecuencia de este truncamiento es que las expansiones variacionales correspondientes dejarán de tener la propiedad de ser extensivas y consistentes en tamaño [5–9].

Un método químico cuántico es consistente en tamaño si la suma de las energías calculadas para dos sistemas A y B es igual a la energía de un supersistema AB en el cual A y B se encuentran separados por una distancia muy grande, es decir, $E(AB) = E(A) + E(B)$.

Por otra parte, un método químico cuántico es extensivo en tamaño si cumple con la condición de que las energía calculadas con dicho método escalan de manera lineal con el número de subsistemas idénticos [7–9]. El ejemplo más común para mostrar ésta propiedad es un sistema compuesto por unidades no interactuantes, los cuales pueden ser de N átomos de He o N moléculas de H_2 . Un método es extensivo en tamaño si $E(N \text{ He}) = N E(\text{He})$.

Sin embargo, estos no son los ejemplos más afortunados, ya que la extensión en tamaño es un concepto mucho más general que la consistencia en tamaño. La extensión en tamaño aplica para cualquier geometría del sistema y no sólo para sistemas no interactuantes, mientras que la consistencia en tamaño aplica únicamente en el caso límite de una separación a distancia infinita de los componentes del sistema[5].

La definición precisa de extensión en tamaño para sistemas arbitrarios es mucho más complicada que la que se muestra en el ejemplo de N átomos de He[9].

En general, se tiende a pensar que la extensión en tamaño y la consistencia en tamaño son propiedades físicas muy importantes y que todos los modelos de la mecánica cuántica deberían incorporar de manera estricta, aunque tal vez éstas características no sean tan importantes [63] o bien ésta falla sea rápidamente compensada al aumentar el tamaño del IC [5].

Los resultados del capítulo anterior muestran que es posible disminuir de manera práctica los errores de consistencia en tamaño que se presentan en los cálculos de IC convencional, al sustituir los orbitales de HF por orbitales naturales producto de cálculos de ICAC. En éste capítulo se presentan algunos resultados que muestran que ICSDTQQnSx proporciona resultados que son consistentes en tamaño y extensivos en tamaño dentro de la exactitud numérica con la cual se llevan a cabo estos cálculos.

7.2. Resultados y discusión.

Para comenzar llevamos a cabo cálculos de ICSDTQQn para un conjunto de sistemas de prueba que contienen cada vez un átomo de He extra, el cual a su vez se encuentra separado de cualquier otro átomo de He por una distancia de al menos 10 \AA . Cada uno de los sistemas de prueba presenta una configuración espacial de simetría molecular perteneciente al grupo d_{2h} o bien a un grupo que contenga a d_{2h} .

En todos los casos se utilizó un conjunto de base cc-pVDZ, el cual es insuficiente para llevar a cabo comparaciones con datos experimentales pero que permite realizar cálculos ICC o con calidad de ICC para todos los átomos aislados, así como para la obtención de todos los valores de ICS reportados.

A pesar de que de antemano se sabe que para los sistemas con hasta 3 átomos de He, ICSDTQQnSx da como resultado valores de la energía libres de errores de consistencia en tamaño, dichos resultados se incluyen en la Tabla 7.1 solo con el fin de llevar a cabo un estudio completo.

# de átomos de He	E(N He)			N E(He)
	ICSD	ICS DTQ	ICS DTQQnSx	
1	-2.887595			-2.887595
2	-5.774726	-5.775190		-5.775190
3	-8.661413	-8.662779	-8.662785	-8.662785
4	-11.547675	-11.550359	-11.550379	-11.550379
5	-14.433529	-14.437925	-14.437974	-14.437974
6	-17.318990	-17.325471	-17.325568	-17.325569
8	-23.088798	-23.100491	-23.100740(20)	-23.100759
10	-28.857205	-28.875386	-28.875890(60)	-28.875948

TABLA 7.1: Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de He una base cc-pVDZ, a excepción de los valores obtenidos para el octámero y decámero, todos los resultados fueron calculados con PSI4.

Los valores de la Tabla 7.1 muestran que hasta $N = 6$ los valores IC hasta séxtuples están “libres del error” de consistencia en tamaño a nivel de 1 microHartree. Para $N = 10$, ICSDTQ presenta un error con respecto al ICC menor a 0.5 miliHartree y el resultado de ICS hasta séxtuples concuerda con el valor “exacto” a nivel de 0.1 miliHartree. El error de truncamiento de ICS de este último resultado es del orden de los 100 miliHartrees debido las limitaciones actuales de ATMOL que no permiten incluir más de 150,000 subclases en el cálculos de ICSP y no necesariamente implican la “falta” de consistencia en tamaño de los resultados de ICSDTQQSx que se obtendrían de manera exacta (aunque el respectivo cálculo debería contener más de 60 mil millones de FSNE). La cota superior de ICS reportada para el IC hasta séxtuples se obtuvo en un espacio que contiene poco más de 2 millones de FSNE.

El siguiente conjunto de prueba fue el de los oligómeros de Li, los resultados se obtuvieron con las mismas coordenadas utilizadas en el estudio de átomos de He.

Los datos en la Tabla 7.2 muestran que los resultados de ICSDTQ difieren en menos de 20 microHartrees respecto a los valores de ICC y los resultados de ICSDTQQnSx difieren en menos de 5 microHartrees.

# de átomos de Li	E(N Li)			N E(Li)
	ICSdT	ICSdTQ	ICSdTQQnSx	
1	-7.432638			-7.432638
2	-14.865221	-14.865277	-14.865277	-14.865275
3	-22.297912	-22.297919	-22.297919	-22.297913
4	-29.729883	-29.730537	-29.730545(10)	-29.730550

TABLA 7.2: Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de Li con una base cc-pVDZ. A excepción de los valores obtenidos para el tetrámero, todos los resultados fueron calculados con PSI4.

# de átomos de Be	E(N Be)			N E(Be)
	ICSdTQ	ICSdTQQn	ICSdTQQnSx	
1	-14.617410			-14.617410
2	-29.234812	-29.234819	-29.234822	-29.234819
3	-43.851211	-43.851243	-43.852225(10)	-43.852229

TABLA 7.3: Energías totales de IC (en Hartree) para el estado fundamental de oligómeros de Be con una base cc-pVDZ. Los valores obtenidos para el monómero y el dímero fueron calculados con PSI4.

# de átomos de B	E(N B)			N E(B)
	ICSdTQ	ICSdTQQn	ICSdTQQnSx	
1	-24.590629	-24.590630		-24.590630
2	-49.181017	-49.181254	-49.181259	-49.181259

TABLA 7.4: Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de B con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.

# de átomos de C	E(N C)			N E(C)
	ICSdTQ	ICSdTQQn	ICSdTQQnSx	
1	-37.761904	-37.761905	-37.761905	-37.761905
2	-75.516647	-75.521316	-75.523731	-75.523809

TABLA 7.5: Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de C con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.

La Tabla 7.3 muestra que en el caso del trímero de Be la energía de IC hasta séxtuples converge al valor de ICC a nivel de 10 microHartrees, un error relativamente pequeño si se toma en cuenta que las excitaciones séxtuples son lo mínimo que se requiere para tomar en cuenta la importante cuasi-degeneración entre $1s^2 2s^2$ y $1s^2 2p^2$ en Be atómico.

# de átomos de N	E(N N)			N E(N)
	ICSdTQ	ICSdTQQ _n	ICSdTQQ _n S _x	
1	-54.480110	-54.480115	-54.480115	-54.480115
2	-108.892089	-108.950391	-108.959410	-108.960230

TABLA 7.6: Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de N con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.

# de átomos de F	E(N F)			N E(F)
	ICSdTQ	ICSdTQQ _n	ICSdTQQ _n S _x	
1	-99.529456	-99.529506	-99.529518	-99.529518
2			-199.058897	-199.059036

TABLA 7.7: Energías totales de IC (en Hartree) para el estado fundamental del monómero y dímero de F con una base cc-pVDZ, todos los resultados fueron calculados con PSI4.

Para continuar con este pequeño estudio se procedió a analizar el error de consistencia en tamaño del resto de los dímeros de los elementos de segundo período (B, C, N, F). La ausencia del dímero de oxígeno se debe a que nuestra versión de PSI4 es incapaz de llevar a cabo el cálculo de ICSdTQQ_nS_x debido al tamaño de la función de onda de IC.

Las Tablas 7.4, 7.5, 7.6 y 7.7 muestran que los errores de truncamiento en tamaño de IC hasta séxtuples es menor al microHartree en el caso del dímero de B, menor a 100 microHartrees en el caso del C, de alrededor de 0.8 miliHartrees (0.5 kcal/mol) en el caso del dímero de N, y menor a 160 microHartrees (0.1 kcal/mol) en el caso del dímero de F. Éstos últimos resultados muestran que el error de consistencia en tamaño no necesariamente esta influido únicamente por la componente de extensión en tamaño, ya que cualquiera esperaría que el error de consistencia en tamaño debería aumentar conforme aumenta el número de partículas (en éste caso electrones) en el sistema, sin embargo, se observa que es más importante el hecho de que ICSdTQQ_nS_x no se desempeña de igual manera al tratar de recuperar energía de correlación para un sistemas con seis electrones desapareados, Esto es, para disociar correctamente, N₂ requiere de excitaciones séxtuples, por lo que un ICSdTQQ_nS_x apenas proporciona una disociación correcta a nivel de modelo de partículas independientes, despreciando la correlación electrónica en lo que concierne a excitaciones de las séxtuples aludidas.

En todos los casos estudiados los resultados de IC hasta séxtuples son consistentes en tamaño dentro de la denominada exactitud química. El caso más singular es el de los sistemas de átomos de He. En el papel, para obtener valores rigurosamente extensivos en tamaño la función de onda de IC debería de incluir excitaciones de orden 20, sin embargo, los resultados obtenidos de las cotas superiores a la energía de ICSDTQQnSx presentan un error con respecto al ICC menor a 100 microHartree, los cuales se recuperan al considerar los errores de truncamiento calculados con la fórmula de Brown.

La explicación de éstos resultados esta relacionado con lo visto en el capítulo 3. En el caso de los sistemas conformados por átomos de He (o por moléculas de H₂), las configuraciones que contribuyen cuantitativamente a la energía del sistema son las configuraciones desconectadas, principalmente aquellas que se forman como productos de configuraciones dobles, es decir, cuádruples, séxtuples, etc.

Es un hecho conocido que la clave para la extensión en tamaño proviene de las configuraciones desconectadas [8]. Esto se confirma al analizar los resultados de ICS en éste tipo de problemas, ya que se observa que la configuraciones conectadas tienen relativamente poca importancia en la composición de la función de onda.

Debido a que la principal fortaleza de ICS es justamente su capacidad para calcular las contribuciones a la energía de las configuraciones desconectadas (usando la fórmula de Brown), ICS presenta una excelente opción para evaluar funciones de onda de sistemas cercanos a la disociación sin la necesidad de llevar a cabo cálculos variacionales enormes, debido a que se puede relizar cálculos con funciones de onda bastante aproximadas y obtener el resto de la energía de correlación mediante la formula de Brown.

La Tabla 7.1 muestra resultados que están en el límite de las capacidades actuales de ATMOL y AUTOCL. El sistema con 10 átomos de Helio con una base cc-pVDZ representa que somos capaces de tratar la correlación electrónica de un sistema de 20 electrones en 50 orbitales, lo cual es complementario al trabajo del capítulo 5 donde se mostraron cálculos de 10 electrones en más de 400 orbitales. Lo anterior es importante, porque muestra una de las posibilidades que se abren para el futuro uso de ATMOL y AUTOCL: trabajar sistemas mucho más grandes utilizando la aproximación de los espacios activos completos.

7.3. Conclusiones

Para un conjunto de sistemas compuestos por átomos idénticos separados a distancia “infinita” obtuvimos cotas superiores a la energía de IC hasta séxtuples que en cada uno de los casos están libres de error de consistencia en tamaño dentro de la exactitud química.

El sistema de estudio más grande está constituido por 10 átomos de He, el valor de ICS hasta séxtuples obtenido, -28.8759 Hartree, presentan un error con respecto al ICC menor a 100 microHartree, los cuales se recuperan al considerar los errores de truncamiento calculados con la fórmula de Brown.

De los dímeros de los elementos del segundo período, el N_2 es el sistema que presenta el mayor error de consistencia en tamaño (alrededor de 0.5 kcal/mol), mientras que en el caso de F_2 este error es menor a 0.1 kcal/mol. Lo cual muestra que existen factores que influyen el error de consistencia en tamaño en mayor medida que el simple número de elementos idénticos en el sistema.

Aunque los resultados obtenidos en éste capítulo no prueban la consistencia en tamaño de IC truncado a sexto orden, si indican que ésta falla de IC sea rápidamente compensada al aumentar el tamaño de la función de onda de IC.

Capítulo 8

Conclusiones

Obtuvimos la mejor cota superior reportada en la literatura a la energía del estado fundamental de la molécula de agua, -76.4343 Hartree, a nivel de ICSDTQQnSx, con una base cc-pCV6Z. Este valor recupera más del 98.8% de la energía de correlación del sistema.

Al llevar a cabo extrapolaciones a CBC, encontrando un valor estimado de la energía no relativista del agua de $-76.4386(9)$ Hartree y al analizar los patrones de convergencia de la contribución a la energía por orden de excitación obtuvimos un valor estimado de $-76.4386(6)$ Hartree.

En el caso de la molécula de H_2O se recupera de manera cuantitativa la contribución a la energía debida a las excitaciones quintuplas y sextuplas usando una base cuádruple zeta.

Para llevar a cabo un ICS más eficiente en el estudio de la disociación simétrica de la molécula de H_2O se puede hacer uso de orbitales naturales de un ICAC, lo cual permite obtener aproximaciones al IC hasta sextuplas que reproducen los valores de ICC con un error menor a 0.1 kcal/mol.

Estos resultados superan en calidad a cualquier método multirreferencial que se encontró en la literatura, lo cual muestra la eficacia de ICS para aproximar los valores de ICC y en la práctica están libres de error de consistencia en tamaño.

Los valores de energía obtenidos dentro de los subespacios S_{01} muestran el rendimiento de ICS cuando se utiliza en una versión aún más simplificada como un modelo de química cuántica.

Para un conjunto de sistemas compuestos por átomos idénticos separados a distancia “infinita” obtuvimos cotas superiores a la energía de IC hasta séxtuples que en cada uno de los casos están libres de error de consistencia en tamaño dentro de la exactitud química.

El sistema de estudio más grande está constituido por 10 átomos de He, el valor de ICS hasta séxtuples obtenido, -28.8759 Hartree, presentan un error con respecto al ICC menor a 100 microHartree, los cuales se recuperan al considerar los errores de truncamiento calculados con la fórmula de Brown.

El sistema con 10 átomos de Helio con una base cc-pVDZ representa que somos capaces de tratar la correlación electrónica de un sistema de 20 electrones en 50 orbitales, lo cual es complementario al trabajo realizado en la obtención de la energía no relativista del H_2O (10 electrones en más de 400 orbitales). Lo cual muestra los límites de las capacidades actuales de nuestros programas, y da una muestra de las posibilidades que se abren para su uso en el futuro: trabajar sistemas mucho más grandes utilizando la aproximación de los espacios activos completos.

Bibliografía

- [1] C.F. Bunge. *J. Chem. Phys.*, 125:014107, 2006.
- [2] C.F. Bunge and R. Carbó-Dorca. *J. Chem. Phys.*, 125:014108, 2006.
- [3] J. Olsen, P. Jorgensen, H. Koch, A. Balkova, and R.J. Bartlett. *J. Chem. Phys.*, 104:8007, 1996.
- [4] T. Shiozaki, M. Kamiya, S. Hirata, and E.F. Valeev. *J. Chem. Phys.*, 130:054101, 2009.
- [5] C.D. Sherrill and H.F. Schaefer III. *Adv. Quant. Chem.*, 34:143–269, 1999.
- [6] I. Shavitt. *The Method of Configuration Interaction*. Methods of Electronic Structure Theory (Modern Theoretical Chemistry), vol. 3, ed. H. F. Schaefer III, Plenum Press, Nueva York, 1977.
- [7] D. Cremer. *WIREs Comput. Mol. Sci.*, 3:482, 2013.
- [8] J.A. Pople. *Theoretical models for chemistry*. in Energy, Structure and Reactivity, ed. D.W. Smith, W.B. McRae, John Wiley & Sons, Nueva York, 1973.
- [9] I. Shavitt and R.J. Bartlett. *Many-Body Methods in Chemistry and Physics. MBPT and Coupled-Cluster Theory*. Cambridge University Press, Nueva York, 2009.
- [10] K. Ruedenberg L. Bytautas. *J. Chem. Phys.*, 124:174304, 2006.
- [11] J. Olsen, P. Jorgensen, and J. Simons. *Chem. Phys. Lett.*, 169:463, 1990.
- [12] C.D. Sherrill and P. Piecuch. *J. Chem. Phys.*, 122:124104, 2005.
- [13] R. Hoffmann, P. von Ragué Schleyer, and H.F. Schaefer III. *Angew. Chem. Int. Ed.*, 47:7164, 2008.

-
- [14] O.L. Polyansky, N.F. Zobov, S. Viti, J. Tennyson, P.F. Bernath, and L. Wallace. *Science*, 277:346, 1997.
- [15] O.L. Polyansky, A.G. Csaszar, S.V. Shirin, N.F. Zobov, P. Barletta, J. Tennyson, D.W. Schwenke, and P.J. Knowles. *Science*, 299:539, 2003.
- [16] G.K.L. Chan and M. Head-Gordon. *J. Chem. Phys.*, 118:8551, 2003.
- [17] P.O. Widmark, P.A. Malmqvist, and B. Roos. *Theo. Chim. Acta*, 77:291, 1990.
- [18] C.X. Almora-Díaz. *J. Chem. Phys.*, 140:184302, 2014.
- [19] J.M. Turney, A.C. Simmonett, R.M. Parrish, E.G. Hohenstein, F.A. Evangelista, J.T. Fermann, B.J. Mintz, L.A. Burns, J.J. Wilke, M.L. Abrams, N.J. Russ, M.L. Leininger, C.L. Janssen, E.T. Seidl, W.D. Allen, H.F. Schaefer, R.A. King, E.F. Valeev, C.D. Sherrill, and T.D. Crawford. *WIREs Comput. Mol. Sci.*, 2:556, 2012.
- [20] R. McWeeny. *Methods of Molecular Quantum Mechanics*. Academic Press, Londres, 2001.
- [21] W. Ritz. *J. Reine Angew. Math.*, 1:135, 1909.
- [22] J.C. Slater. *Phys. Rev.*, 34:1293, 1929.
- [23] P. Jönsson C. Froese-Fischer, T. Brage. *Computational Atomic Structure: An MCHF Approach*. Institute of Physics Publishing, Bristol, 1997.
- [24] H.F. Schaefer III. *The Electronic Structure of Atoms and Molecules: A Survey of Rigorous Quantum Mechanical Results*. Adisson-Wesley, Reading, MA, 1972.
- [25] N.M. Gray and L.A. Willis. *Phys. Rev.*, 38:248, 1931.
- [26] P.-O. Löwdin. *Phys. Rev.*, 97:1509, 1955.
- [27] P.-O. Löwdin. *Adv. Phys.*, 5:1, 1956.
- [28] P.-O. Löwdin. *Rev. Mod. Phys.*, 34:520, 1962.
- [29] P.-O. Löwdin. *Rev. Mod. Phys.*, 36:966, 1964.
- [30] E. Davidson. *Rev. Mod. Phys.*, 44:451, 1972.
- [31] C.F. Bunge. *Theor. Chem. Acta*, 16:126, 1970.

- [32] G.A. Pettersson, D.K. Mallick, M.J. Frisch, and M. Braunstein. *J. Chem. Phys.*, 123:074111, 2005.
- [33] T.H. Dunning. *J. Chem. Phys.*, 90 (2):1007, 1989.
- [34] M.W. Schmidt and M.S. Gordon. *Annu. Rev. Phys. Chem.*, 49:233, 1998.
- [35] S. Hirata and R.J. Bartlett. *Chem. Phys. Lett.*, 321:216, 2000.
- [36] A. Bunge. *J. Chem. Phys.*, 53:20, 1970.
- [37] C.F. Bunge and A. Bunge. *Int. J. Quantum Chem.*, 7:927, 1973.
- [38] B. Huron, J.P. Malrieu, and P. Rancurel. *J. Chem. Phys.*, 58:5745, 1973.
- [39] S. Evangelisti, J.P. Daudey, and J.P. Malrieu. *Chem. Phys.*, 75:91, 1983.
- [40] R. Brown. *A Configuration Interaction Study of the States of the LiH molecule*. PhD thesis, Departamento de Química, Indiana University, 1967.
- [41] C.F. Bunge. *Phys Rev. A*, 14:1965, 1976.
- [42] O. Sinanoğlu. *J. Chem. Phys.*, 36:706, 1962.
- [43] F. Coester. *Nucl. Phys.*, 7:421, 1958.
- [44] C.F. Bunge. *Theor. Chem. Acc.*, 126:139, 2009.
- [45] E.R. Davidson. *J. Comput. Phys.*, 15:87, 1975.
- [46] E.R. Davidson. *Comput. Phys.*, 7:519, 1994.
- [47] M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. van Dam, D. Wang, J. Nieplocha, E. Apra, T.L. Windus, and W.A. de Jong. *Comput. Phys. Commun.*, 181:1477, 2010.
- [48] T. Helgaker, W. Klopper, H. Koch, and J. Noga. *J. Chem. Phys.*, 106:9639, 1997.
- [49] A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen, and A.K. Wilson. *Chem. Phys. Lett.*, 286:243, 1998.
- [50] W. Klopper, K.L. Bak, P. Jorgensen, and J. Olsen T. Helgaker. *J. Phys. B*, 32:R103, 1999.
- [51] A. Galano J.R. Álvarez Idaboy. *Theor. Chem. Acc.*, 126:75, 2010.

- [52] K. Hirao. *Recent Advances in Multireference Methods*. in *Recent Advances in Computational Chemistry*, vol. 4, ed. M.R. Hoffmann, K.G. Dyall, World Scientific, Singapur, 1999.
- [53] K.G. Dyall M.R. Hoffmann. *Low-Lying Potential Energy Surfaces*. in *ACS Symposium Series*, vol. 828, ed. M.R. Hoffmann, K.G. Dyall, American Chemical Society, Washington, 2002.
- [54] M.L. Abrams and C.D. Sherrill. *Chem. Phys. Lett.*, 412:121, 2005.
- [55] L. Fusti-Molnar and P.G. Szalay. *J. Phys. Chem.*, 100:6288, 1996.
- [56] M. Kállay, P.G. Szalay, and P.R. Surján. *J. Chem. Phys.*, 117:980, 2002.
- [57] P.G. Szalay, T. Muller, G. Gidofalvi, H. Lischka, and R. Shepard. *Chem. Rev.*, 112:108, 2012.
- [58] S. Canuto B.J. Costa-Cabrera. *Chem. Phys. Letts.*, 406:300, 2005.
- [59] J. Paldus, M. Takahashi, and B.W.H. Cho. *Int. J. Quantum Chem.*, S18:237, 1984.
- [60] M. Head-Gordon, T.V. Voochris, S.R. Gwaltney, and E.F.C. Byrd. *Low-Lying Potential Energy Surfaces*. in *ACS Symposium Series*, vol. 828, ed. M.R. Hoffmann, K.G. Dyall, American Chemical Society, Washington, 2002, p. 93.
- [61] J. Paldus X. Li. *Low-Lying Potential Energy Surfaces*. in *ACS Symposium Series*, vol. 828, ed. M.R. Hoffmann, K.G. Dyall, American Chemical Society, Washington, 2002, p. 10.
- [62] P. Piecuch, K. Kowalski, I.S.O. Pimienta, and S.A. Kucharski. *Low-Lying Potential Energy Surfaces*. in *ACS Symposium Series*, vol. 828, ed. M.R. Hoffmann, K.G. Dyall, American Chemical Society, Washington, 2002, p. 31.
- [63] W. Duch and G.H.F. Diercksen. *J. Chem. Phys.*, 101:3018, 1994.

Highly correlated configuration interaction calculations on water with large orbital bases

César X. Almora-Díaz^{a)}

Instituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, México 01000, México

(Received 8 January 2014; accepted 21 April 2014; published online 8 May 2014)

A priori selected configuration interaction (SCI) with truncation energy error [C. F. Bunge, J. Chem. Phys. **125**, 014107 (2006)] and CI by parts [C. F. Bunge and R. Carbó-Dorca, J. Chem. Phys. **125**, 014108 (2006)] are used to approximate the total nonrelativistic electronic ground state energy of water at fixed experimental geometry with CI up to sextuple excitations. Correlation-consistent polarized core-valence basis sets (cc-pCVnZ) up to sextuple zeta and augmented correlation-consistent polarized core-valence basis sets (aug-cc-pCVnZ) up to quintuple zeta quality are employed. Truncation energy errors range between less than 1 μ hartree, and 100 μ hartree for the largest orbital set. Coupled cluster CCSD and CCSD(T) calculations are also obtained for comparison. Our best upper bound, -76.4343 hartree, obtained by SCI with up to sextuple excitations with a cc-pCV6Z basis recovers more than 98.8% of the correlation energy of the system, and it is only about 3 kcal/mol above the “experimental” value. Despite that the present energy upper bounds are far below all previous ones, comparatively large dispersion errors in the determination of the extrapolated energies to the complete basis set do not allow to determine a reliable estimation of the full CI energy with an accuracy better than 0.6 mhartree (0.4 kcal/mol). © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4874319>]

I. INTRODUCTION

Configuration interaction (CI) is a straightforward and general approach for the treatment of electron correlation in atoms and molecules.¹ In particular, full configuration interaction (FCI) gives the exact solution to the Schrödinger equation within the chosen basis, and it is of interest in connection with benchmarks for calibrating results using approximate wave functions other than FCI. Nevertheless, these have only been attempted with rather short orbital bases, as the total number of configurations easily grows beyond manageable size.

Recently, *a priori* selected CI (SCI)² and CI by parts (CIBP)³ have been proven to be useful alternatives to approximate CI energies using up to decuple excitations in atomic systems.⁴ One of the features of SCI is the possibility to obtain numerical solutions of the Schrödinger equation with controlled and hence predictive accuracy, which is one of the aims of electronic structure theory.⁵

In this paper, the first implementation of SCI and CIBP for molecular systems is reported. This new functionality is incorporated in the latest versions of the electronic structure programs ATMOL and AUTOCL.⁴ In this regard, it is necessary to mention that all molecular integrals over primitive basis functions in this paper are obtained from a slightly modified version of the program NWChem.⁶ Program ATMOL then performs the four-index transformation after which molecular integrals over symmetry orbitals are extracted for use in (symmetry) natural orbital evaluation, SCI and CIBP.

In order to exhibit the accuracy of our work, we calculate nonrelativistic ground state energies of the water molecule at its experimental equilibrium geometry reproducing truncated all electron highly correlated CI energies: CI up to triple excitations (CISDT), CI up to quadruple excitations (CISDTQ), CI up to quintuple excitations (CISDTQQn), and CI up to sextuple excitations (CISDTQQnSx), with correlation consistent basis sets that include correlating functions for the core electrons, cc-pCVnZ (with $n = D, T, Q, 5, 6$) and aug-cc-pCVnZ (with $n = D, T, Q, 5$) with further extrapolation to the complete basis set limit.

We choose H₂O ground state because it has become a benchmark system for full *ab initio* quantum mechanical treatments.^{7,8} Although we would have liked to perform all calculations with the aug-cc-pCVnZ basis sets, unfortunately we have not been able to carry out calculations with the aug-cc-pCV6Z orbital set because, with the present version of our program, we need more than 350 GB of RAM to perform the 4-index transformation of its 533 primitive functions (more information in connection with the results). Therefore, our study is primarily based on the results with the cc-pCVnZ bases. The largest orbital set used, cc-pCV6Z, includes 412 primitive functions.

Truncation CI energy errors range between less than 1 μ hartree for all calculations with the cc-pCVDZ orbital set, to less than 100 μ hartree (0.06 kcal/mol) for the cc-pCV6Z.

Based on the results for double zeta basis sets^{8,9} and up to quadruple zeta basis sets¹⁰ in water, and on the results for the ten electron Ne ground state,⁴ the correlation energy from septuples up to decuples are expected to be considerably smaller than the above error, so we decided not to incorporate septuples onwards. However, in the light of the foregoing

^{a)}xalmora@fisica.unam.mx. In partial fulfillment for the requirements of the degree of Doctor in Chemistry at the Universidad Nacional Autónoma de México.

we note that our CISDTQnSx results are as good as FCI to within an error less than 160 μ hartree (0.1 kcal/mol).

Our approach differs for the best FCI approximation to date, made by Bytautas and Ruedenberg,¹⁰ who did an excellent job to calculate the full valence CI using cc-pVnZ basis sets with up to quadruple zeta, and then extrapolated to the complete basis set limit using previous results of CCSD(T) and taking the core valence correlation energy from CCSD(T)-R12 calculated from Klopper¹¹ as well as from Noga and Klopper.¹² Instead, we correlate all ten electrons and estimate truncation energy errors of discarded configurations after examining all configurations provided by the orbital basis.

The calculations were performed at the experimental equilibrium geometry given by Helgaker *et al.*¹³ For comparison purposes, we also carried out calculations at the level of coupled cluster theory with singles and doubles excitations (CCSD), and singles and doubles excitations with a perturbative triples correction (CCSD(T)).

Various attempts to adjust the calculated data in order to obtain reliable values for the extrapolated complete basis set energies were carried out, however, we were unable to find a formula yielding results of precision comparable to our SCI treatments.

II. SELECTED CI

Since SCI and CIBP are well described in Refs. 2–4, only the essentials are outlined here. Schrödinger's equation can be approximated by a matrix-eigenvalue equation

$$\mathbf{H}\mathbf{C}_\mu = E_\mu^{\text{FCI}}\mathbf{C}_\mu, \quad (1)$$

where \mathbf{H} is the representation of H in terms of Slater determinants or N-electron symmetry-eigenfunctions constructed from a given orbital basis. Equation (1) defines the FCI method¹⁴ and E_μ^{FCI} is the FCI energy. In terms of FCI quantities, the exact eigenvalues E_μ of Schrödinger's equation may be expressed as

$$E_\mu = E_\mu^{\text{FCI}} + \Delta E_\mu^{\text{BSIE}}, \quad (2)$$

where $\Delta E_\mu^{\text{BSIE}}$ is the basis set incompleteness error.^{15–18} Henceforth, the subscript μ will be dropped.

The CI expansion is written in terms of symmetry-eigenfunctions F_{gK} ^{19,20} as

$$\Psi = \sum_{K=1}^{K_x} \sum_{g=1}^{g_K} F_{gK} C_{gK}, \quad (3)$$

$$F_{gK} = O(\Gamma, \gamma) \sum_{i=1}^g D_{iK} b_i^g = \sum_{i=1}^{n_K} D_{iK} c_i^g, \quad g = 1, \dots, g_K. \quad (4)$$

K and g label configurations and degenerate elements, respectively. F_{gK} is a configuration-state-function (CSF) expressed as a linear combination of n_K Slater determinants D_{iK} . $O(\Gamma, \gamma)$ is a symmetric projection operator²¹ for all pertinent symmetry operators Γ and a given (N-electron) irreducible representation γ .^{22–25}

Normalized symmetry-configurations, Φ_K , are obtained by summing over the g index in (3) and renormalizing

$$\Phi_K = N_K \sum_{g=1}^{g_K} F_{gK} C_{gK}, \quad (5)$$

therefore, in terms of Φ_K the CI expansion is rewritten as

$$\Psi = \sum_{K=1}^{K_x} \Phi_K B_K, \quad (6)$$

$$B_K = \frac{1}{N_K}, \quad N_K = \sqrt{1 / \sum_{g=1}^{g_K} C_{gK}^2}, \quad (7)$$

for the only purpose to select configurations: actual CI calculations are still carried out via Eq. (3).

Configurations are of two types: disconnected ones, when the determinants can be written as antisymmetrized Hartree products of lower excited determinants, care being taken of dividing by an appropriate number of reference Hartree products. All others K 's are called connected ones, with their determinantal CI coefficients $c_{ijk\dots}^{abc\dots}$ being related² to the corresponding amplitudes $t_{ijk\dots}^{abc\dots}$ of coupled cluster theory.^{26,27}

The SCI method² seeks to approximate FCI (or truncated CI) results by performing CI on quantitatively *a priori* selected spaces while *a priori* estimating truncation energy errors for the discarded spaces.

Selection of configurations is made using the Brown formula²⁸

$$\Delta E_K = (E - H_{K,K})C_K^2 / (1 - C_K^2), \quad (8)$$

where E is any pertinent energy, and the diagonal matrix element $H_{K,K}$ can be well approximated by the expectation value of any determinant of K .² In this way, the values of g_K and n_K are eliminated. The strength of SCI stems from the possibility to predict B_K 's fairly well for disconnected configurations. As mentioned in Ref. 2 after Eq. (25), reliable prediction of B_K coefficients requires complicated formulae whose FORTRAN implementation was achieved through a special program that produced FORTRAN subroutines both for formula evaluation and to find out whether a given configuration is a disconnected one or not.

For a connected q-excited configuration K , the energy contribution ΔE_K^{con} can be approximated by²⁹

$$\Delta E_K^{\text{con}} = (E - H_{K,K})P(q, K), \quad (9)$$

$$P(q, K) = \prod_{i=1}^q n_{K_i}, \quad (10)$$

where n_{K_i} is the occupation number of the correlation natural orbital K_i . Having an energy dimension, ΔE_K^{con} is called a pseudo-energy just to keep in mind that its accuracy may be rather poor, nevertheless, it provides a surprisingly efficient selection criterion when Brown's formula is inapplicable, viz., for connected configurations.

In atomic calculations, the first step of SCI is developing an orbital basis. In molecular calculations aiming to compare

results with those obtained by other methods, however, the use of standard and well known orbital bases is preferable.

The next step is to reduce the size of the FCI space by truncating the list of configurations that will be spanning a model space

$$E^{\text{FCI}} = E_M + \Delta E_M^{\text{egy}} + \Delta E_M^{\text{psegy}}, \quad (11)$$

where E is the energy in the model space, ΔE_M^{egy} is the total truncation energy error calculated by Brown's formula, and $\Delta E_M^{\text{psegy}}$ is the total truncation pseudo-energy error calculated by the modified Brown formula, Eq. (10). Reduction of the FCI space to the model space is carried out by program A TOC⁴ to evaluate lists of symmetry eigenfunctions without invoking matrix-eigenvalue equations: the only needed matrix elements are those required to evaluate ΔE_M^{egy} and $\Delta E_M^{\text{psegy}}$.

So far, the only sure thing about $\Delta E_M^{\text{psegy}}$ is that the magnitude of this error is stable around $1 \mu\text{hartree}$ when the $\Delta E_M^{\text{psegy}}$ is around $0.0003 \mu\text{hartree}$.

On the other hand, ΔE_M^{egy} is a considerably more reliable quantity:² if ΔE_M^{egy} is of the order of a few $\mu\text{hartree}$ it may be verified that this result is reliable to within a fraction of one $\mu\text{hartree}$, whereas if it is $100 \mu\text{hartree}$, the true error may be a smaller one, as much as one half the estimated value.

For very large orbital bases, particularly in absence of symmetry, the orbitals involved in n -excited configurations must be truncated according to thresholds for natural orbital occupation numbers depending on n , and a corresponding sensitivity analysis must be carried out or else the truncation error $\Delta E_M^{\text{egy}} + \Delta E_M^{\text{psegy}}$ remains unknown. However, the size of the basis sets used in this work allows to include all symmetry orbitals to produce up to sextuple-excited configurations.

In atomic calculations, the main stumbling block is the occurrence of configurations with very large degeneracies, beyond several thousands. In molecules, the absence of orbital angular momentum symmetry causes degeneracies to be comparatively small: it takes only 3 s to generate all needed C_{2v} eigenfunctions up to sextuples. Moreover, most configurations in (6) are disconnected ones, thus the selection process becomes computationally less demanding and more accurate than for atoms.

After obtaining a model space without invoking matrix-eigenvalue equations, a corresponding configuration list of symmetry eigenfunctions is produced, again a rather trivial step in molecules compared with the huge number of determinants appearing in atoms together with large degeneracies. Together with this list, we get the corresponding truncation energy error $\Delta E_M^{\text{egy}} + \Delta E_M^{\text{psegy}}$.

The model space is still too large for a straight CI approach and needs to be further truncated into a space obtained by invoking different energy and pseudo-energy selection thresholds for each excitation order of disconnected configurations, and different pseudo-energy selection thresholds for each excitation order of connected configurations. The need to apply pseudo-energy selection thresholds to formally disconnected configurations arises by the fact that many of them have a significant connected part, viz., the disconnected terms are made up by both disconnected and connected contributions.

TAB E I. Total energies (in hartree) for H₂O ground state in a TZ2P ANO orbital basis set.

ethod	Chan and Head-ordon ⁸	This work
CISDTQ	-76.314001	-76.314001
CISDTQQn		-76.314511(1)
CCSDTQ	-76.314696	
CISDTQQnSx		-76.314701(5)
CISDTQQnSxSp		-76.314709(5)
D R	-76.31471	
CISDTQQnSxSpOc		-76.314711(5)
FCI	-76.314715(10)	-76.314711(5)

Thus, the model space is further reduced to a selected space

$$E_M = E_S + \Delta E_S^{\text{egy}} + \Delta E_S^{\text{psegy}}, \quad (12)$$

where the meanings ΔE_S^{egy} and $\Delta E_S^{\text{psegy}}$ are identical as those in Eq. (11) having the subscript but now they refer to the truncation of the model space into the selected configuration space. As the selection thresholds become smaller, better energies are obtained with smaller truncation energy errors.

III. CI TS

The configuration list described in Sec. II defines a model space out of which a selected space is obtained. Usually is still too expensive-to-evaluate and extremely huge in size, making a direct CI too inefficient. Thus, a CI by parts procedure is in order.³

The idea is to split in subspaces $S_0, S_1, S_2, \dots, S_r$, of dimensions $d_0, d_1, d_2, \dots, d_r$, respectively. S_0 is a reference space where CI coefficients are always free to vary, and all other subspaces $S_i, i = 1, 2, \dots, r$ will be taken up variationally one after the other.³ The final eigenvalue, E^B , will satisfy

$$E_S = E^{CIBP} + \Delta E^{CIBP}, \quad (13)$$

where ΔE^B is necessarily a positive quantity.

In order to have ΔE^B rather small, the energy corresponding to S_0+S_1 must be as close as possible to E^B , thus pushing up values of d_0+d_1 that require solving the eigenproblem outside RA. After S_1 is processed, however, the remaining subspaces S_i (i greater than 1) are (automatically) chosen so that all matrix elements in $d_0+(i-1)d_i$ fit in RA.

TAB E II. SCF energies for H₂O ground state in hartree.

Basis set (n)	SCF/RHF	Basis set (n)	SCF/RHF
cc-pCVDZ	-76.027204	aug-cc-pCVDZ	-76.041725
cc-pCVTZ	-76.057358	aug-cc-pCVTZ	-76.060839
cc-pCVQZ	-76.064948	aug-cc-pCVQZ	-76.066094
cc-pCV5Z	-76.067105	aug-cc-pCV5Z	-76.067330
cc-pCV6Z	-76.067411	aug-cc-pCV6Z	-76.067466
Fitting Eq. (14)			
Estimated limit (3-6)	-76.0677		-76.0676
Estimated limit (4-6)	-76.0675		-76.0675

TAB E III. Correlation energies for H₂O ground state in hartree. The corresponding upperbounds can be calculated by adding the respective SCF energy.

Basis set (n)	CISD	CISDT	CISDTQ	CISDTQQn	CISDTQQnSx	CCSD	CCSD(T)
cc-pCVDZ	-0.241765	-0.244849	-0.255271	-0.255484	-0.255661	-0.251750	-0.255038
cc-pCVTZ	-0.310648	-0.317919	-0.33203	-0.3326	-0.3328	-0.324186	-0.332421
cc-pCVQZ	-0.331820	-0.340055	-0.3554	-0.3560	-0.3562	-0.346496	-0.356018
cc-pCV5Z	-0.338901	-0.347465	-0.3631	-0.3637	-0.3640	-0.353947	-0.363944
cc-pCV6Z	-0.341671	-0.35034	-0.3660	-0.3667	-0.3669	-0.356832	-0.366990
Fitting Eq. (18)							
Estimated limit (3-6)	-0.3454	-0.3542	-0.3700	-0.3706	-0.3709	-0.3607	-0.3712
Estimated limit (4-6)	-0.3448	-0.3535	-0.3691	-0.3697	-0.3700	-0.3600	-0.3704

I. ELECTRONIC STRUCTURE DISCUSSION

A. INTRODUCTION

All the calculations in this study were carried out at the experimental equilibrium geometry of water, $R_{\text{O-H}} = 95.72$ pm, and $\angle_{\text{H-O-H}} = 104.52^\circ$ with full group symmetry C_{2v} .

CISD, CCSD, and CCSD(T) calculations were carried out using N Chem. SCI calculation to approximate the corresponding energy values of CISDT, CISDTQ, CISDTQQn, and CISDTQQnSx were performed using our programs A TOC and AT O coupled to a modified version of N Chem program,⁶ from which we obtain SCF wave functions and lists of one- and two-electron integrals over a basis of primitive orbitals as input to AT O .

We used the standard aug-cc-pCVnZ and cc-pCVnZ basis set as given in the data basis of N Chem, that is the same one that can be obtained from the E S web page (<https://bse.pnl.gov/bse/portal>).

B. SCF CALCULATION

In order to validate program AT O , we first reproduced two well known benchmarks for the electronic ground state of water molecule. First, we reproduce all CI values reported by Olsen *et al.*,⁹ at equilibrium distance, using the cc-pVDZ (24 orbitals) orbital set, with a truncation energy error less than $1 \mu\text{hartree}$. Then, we reproduced CISD, CISDT, CISDTQ energies reported by Chan and Head-ordon,⁸ using a triple-zeta, double polarization (TZ2P) quality basis (41 orbitals). Furthermore, we calculate other highly correlated CI energies reported in Table I.

Chan and Head-ordon⁸ reported as best upper bound a value of -76.31471 hartree, and an extrapolated value of FCI of -76.314715 ± 0.000010 hartree. We report an upper bound of -76.314711 with a truncation error of

$5 \mu\text{hartree}$ in excellent agreement with the result of Chan and Head-ordon.⁸

Our best upper bound with the TZ2P basis set was carried out using a selected model space of 18.5×10^6 CSFs obtained from a total space larger than 9×10^9 CSFs. This calculation took less than 12 h in a node using 16 processors on E5-2670, 64 GB of RAM and less than 160 GB of disk.

For the double-zeta basis sets, the reported errors are calculated using Brown's formula. For large databases, the errors reported are estimated based on sensitivity analyses which mainly involve the size of the corresponding S_0+S_1 subspace. All energy and pseudoenergy thresholds are set sufficiently low to guarantee a truncation energy error less than $10 \mu\text{hartree}$, thus the main source of error should be ΔE^B , that in all cases we estimate to be less than $100 \mu\text{hartree}$. In this regard, in the case of the ground state of water molecule, the contribution to the energy of the connected configurations is small, less than $10 \mu\text{hartree}$ in the case of the cc-pCVDZ basis set. However, as their number is relatively small, a large portion of them has been introduced via the pseudoenergy thresholds, in order to assure the desired accuracy.

From Table I, we can see that CISDTQQnSx reproduces FCI energies with an error less than $10 \mu\text{hartree}$. This result, together with FCI results on Ne atom where septuples to decuples contribute less than $30 \mu\text{hartree}$ to the total energy, allow us to say that our CISDTQQnSx results reproduce the FCI with an error less than $160 \mu\text{hartree}$, the main source of error being ΔE^B in Eq. (13). This error will likely be made negligibly small after using results from *a priori* SCI to launch an *a posteriori* SCI using Brown's formula and the obtained *a priori* SCI expansion coefficients κ .

C. CONCLUSIONS

For the cc-pVnZ and cc-pCVnZ basis sets families, the Hartree-Fock energies of the n-tuple-zeta basis sets approach

TAB E IV. Correlation energies for H₂O ground state in hartree. The corresponding upperbounds can be calculated by adding the respective SCF energy.

Basis set (n)	CISD	CISDT	CISDTQ	CISDTQQn	CISDTQQnSx	CCSD	CCSD(T)
aug-cc-pCVDZ	-0.256091	-0.261058	-0.273341	-0.27374	-0.27397	-0.267812	-0.273298
aug-cc-pCVTZ	-0.315388	-0.32341	-0.3383	-0.3389	-0.3391	-0.329593	-0.338797
aug-cc-pCVQZ	-0.333710	-0.34220	-0.3578	-0.3584	-0.3587	-0.348645	-0.358541
aug-cc-pCV5Z	-0.339688	-0.34832	-0.3640	-0.3646	-0.3649	-0.354810	-0.364936

TAB E V. Energy contribution by excitation level in hartree.

Basis set (n)	CISD	CISDT	CISDTQ	CISDTQQn	CISDTQQnSx
cc-pCVDZ	-0.241765	-0.00308	-0.01042	-0.00021	-0.00018
cc-pCVTZ	-0.310648	-0.00726	-0.01412	-0.00054	-0.00024
cc-pCVQZ	-0.331820	-0.00824	-0.01532	-0.00060	-0.00026
cc-pCV5Z	-0.338901	-0.00856	-0.01564	-0.00062	-0.00027
cc-pCV6Z	-0.341671	-0.00867	-0.01570	-0.00063	-0.00027
Estimated limit	-0.3454	-0.0088	-0.0157	-0.0006	-0.0003

CBS limits exponentially,^{10,13,30}

$$E_{HF}(X) = E_{HF}(CBS) + a \exp -bX. \quad (14)$$

Equation (14) has proven to be an effective procedure to estimate a near-exact limit for the SCF-Hartree-Fock energy. In our case, this equation gives us the best fit to the $E_{HF}(B)$ energy. Table II contains our calculated SCF energies and the extrapolated SCF-RHF energy value for the cc-pCVnZ and the aug-cc-pCVnZ basis sets. One can see that all results are around the value of -76.0676 hartree.

D. C S

The correlation energy has been found to converge to its CBS limit more slowly than SCF, namely, close to an inverse third power law^{10,13,30}

$$E_{corr}(X) = E_{corr}(B) + aX^{-3}. \quad (15)$$

Other well-known formula used for atomic calculations studied was

$$E_{corr}(X) = E_{corr}(B) + a(X+b)^{-3}, \quad (16)$$

which in turn is a special case of the more general expression

$$E_{corr}(X) = E_{corr}(B) + a(X+b)^{-3+c}, \quad (17)$$

that has been used by Helgaker *et al.*¹³

However, we found a best fit using another widely used formula^{17,32}

$$E_{corr}(X) = E_{corr}(B) + a(X)^{-3} + b(X)^{-4}. \quad (18)$$

Tables III and IV contain our calculated highly correlated energies for both coupled cluster and configuration interaction methods for the cc-pCVnZ and the aug-cc-pCVnZ basis set, respectively. We notice that as the size of the basis set increases, the values of CCSD(T) move below our CISDTQ values, which might indicate that as the basis size increases, CCSD(T) overestimates the correlation energy.

Fitting data to Eqs. (16) or (18) gave us the best results, in the sense that they best reproduce the calculated values. In both cases, fitting data from $n = 3, 4, 5, 6$, or $n = 4, 5, 6$, reproduce calculated values with four decimals (of hartree) and their respective CBS values differ in less than 20 μ hartree. So we decide to use only Eq. (18) to carry out our study.

Several studies^{10,13,17,30,32} indicate that the best fit is performed with values of n greater than 3 and that, wherever possible, an orbital basis set with $n = 6$ should be included within the fitted data set, we exclude all the results with the cc-pCVDZ basis set, and we do not perform the CBS extrapolations with the aug-cc-pCVnZ bases, as yet we cannot get results with the aug-cc-pCV6Z basis. Table III contains our calculated correlation energies for each CI-level and their respective extrapolated values.

One can observe that the difference between fitting with $n = 3-6$ and with $n = 4-6$ gradually increases from 0.6 for the CISD to 0.9 mhartree for the CISDTQ and this difference remains constant up to CISDTQQnSx. In this way, our estimate for the FCI energy of the water with the cc-pCVnZ basis sets is

$$\begin{aligned} & -76.0677(\text{Fit}) - 0.3709(\text{DnBlit}) \\ & = -76.4386(9) \text{ hartree}, \end{aligned} \quad (19)$$

which is in accord with the value derived from experiment, -76.4389, and with the extrapolated value of -76.4390(4) obtained by Bytautas and Ruedenberg.¹⁰

E. C

In Table V, we show the energy contribution to the correlation energy at each excitation level calculated as follows: for CISD is equal to the difference between the CISD energy and the SCF energy, for CISDT is equal to $E(\text{CISDT}) - E(\text{CISD})$, and so on.

From Table V, we can see that in the case of CISDTQ, CISDTQQn, and CISDTQQnSx the energy contributions

TAB E VI. Basis set and CI sizes used in this work.

Basis set (n)	functions	CISD	CISDT	CISDTQ	CISDTQQn	CISDTQQnSx
cc-pCVDZ	28	1917	47 758	691 182	3 625 640	1 065 155
cc-pCVTZ	71	14 699	1 096 867	27 412 581	64 372 755	163 685 771
cc-pCVQZ	144	63 336	9 431 329	50 583 371	102 219 980	226 552 425
cc-pCV5Z	255	201 801	17 405 027	73 671 958	133 976 227	262 575 237
cc-pCV6Z	412	530 077	18 019 330	100 185 900	164 155 506	292 078 001

TAB E VII. Basis set and CI sizes used in this work.

Basis set (n)	functions	CISD	CISDT	CISDTQ	CISDTQQn	CISDTQQnSx
aug-cc-pCVDZ	45	5608	247 466	3 880 829	16 219 136	41 055 213
aug-cc-pCVTZ	105	33253	2 694 662	29 298 868	80 979 764	183 635 656
aug-cc-pCVQZ	201	124 890	11 030 844	60 115 073	125 813 254	253 869 286
aug-cc-pCV5Z	341	362 620	23 550 037	82 432 987	148 960 984	277 845 967

clearly converge to a well defined value at the 0.1 mhartree level, and for CISDT it is almost the case. The only case that does not reach convergence is CISD.

At the bottom of Table V, we show the estimated limits to the energy at each excitation level, for CISD we put the CBS extrapolated with $n = 3-6$, for CISDT we perform a similar calculation, and for CISDTQ, CISDTQQn, and CISDTQQnSx we show the converged values.

Now we are able to estimate the FCI energy by adding to the extrapolated CBS values of SCF, CISD plus the extrapolated contribution of CISDT to the converged energy contributions of CISDTQ, CISDTQQn, and CISDTQQnSx

$$\begin{aligned}
 E(F) &= -76.0677(B - F) - 0.3454(B - D) \\
 &\quad - 0.0088(e \text{ trapolate} - D) \\
 &\quad - 0.0157(\text{on erge} - D) \\
 &\quad - 0.0006(\text{on erge} - D_n) \\
 &\quad - 0.0003(\text{on erge} - D_{n'}) \\
 &= -76.4385(6) \text{ hartree.} \quad (20)
 \end{aligned}$$

The reported error is the one associated to the CBS-CISD extrapolation. This result agrees with the energy value from the pure CBS extrapolation, and also we reduce the uncertainty associated with the inclusion of the quadruple, quintuple, and sextuples excitations.

C

In Tables VI and VII, we show CI sizes used to obtain the corresponding upperbounds showed in Tables III and IV. For the cc-pCV6Z basis set, we can see that the number of CSF representing quintuple and sextuples excitations changes very little with respect to the cc-pCV5Z basis sets.

With the cc-pCV6Z basis the CISDTQ calculation took 156 h, CISDTQQn 225 h, and CISDTQQnSx 352 h using a single

TAB E VIII. Estimated total energies for H₂O ground state in hartree.

uchow <i>et al.</i> , ³¹ iCAS-CI/aug-cc-pCV5Z	-76.4274 (upper bound)
uchow <i>et al.</i> , ³¹ FCI + CBS	-76.431 ± 0.003
This work, CISDTQQnSx/cc-pCV6Z	-76.4343 (upper bound)
Shiozaki <i>et al.</i> , ⁵ CCSDTQ-R12	-76.437 ± 0.003
This work, FCI + CBS(cc-pCVnZ)	-76.4386 ± 0.0009
This work, FCI + CBS(cc-pCVnZ) + converge patterns	-76.4385 ± 0.0006
Halkier <i>et al.</i> , ³⁰ CCSD(T) + CBS	-76.4386
Helgaker <i>et al.</i> , ¹³ CCSD(T)/R12	-76.439 ± 0.002
Bytautas and Ruedenberg, ¹⁰ FCI + CBS	-76.4390 ± 0.0004

node with 16 processors on E5-2670, 64 B of RAM and less than 800 B of disk.

As far as we know, the results of CISDTQ, CISDTQQn, and CISDTQQnSx with the cc-pCV6Z basis (-76.4334, -76.4341, -76.4343 hartree, respectively) represent the best upper bounds to the total energy for the water molecule to date. The lowest upper bound that we found in the literature is -76.4274 hartree from chow *et al.*³¹ The CISDTQQnSx result recovers more than 98.8% of the correlation energy of the system, and it is only about 3 kcal/mol above the experimental value, -76.4389 hartree.¹⁰

Finally, in Table VIII we show our best theoretical estimates of the total and correlation energies of H₂O at equilibrium geometry and compare with other theoretical results.

CONCLUSIONS

The best CI results to date for the ground state of the water molecule with the cc-pVnZ and aug-cc-pCVnZ basis sets have been presented. The CISDTQQnSx results reproduce FCI values within each orbital set to within an error less than 160 μhartree. Furthermore, we present the best upper bound to the nonrelativistic total energy for the ground state of water molecule, -76.4343 hartree, obtained by SCI with up to sextuple excitations with a cc-pCV6Z basis. This result recovers more than 98.8% of the correlation energy of the system, and it is only about 3 kcal/mol above the experimental value.

With our results we estimated CBS for each CI excitation level up to CISDTQQn and with those estimate values we compute a FCI energy of -76.4386(9) hartree. Along with this, analyzing the patterns of convergence of the contributions to the energy at each excitation level up to sextuples, we calculated an estimated value of the FCI of -76.4385(6) hartree.

Programs ATOM and ATOC are currently underlying a major generalization and optimization process together with incorporation of new SCI and CIBP methodologies aiming at significant increases in accuracy and efficiency. As soon as these new program versions become available, we shall proceed to carry out aug-cc-pCV6Z calculations together with CBS extrapolations.

ACKNOWLEDGMENTS

I am indebted to my supervisor, Professor Carlos F. Bunge, for his continued help, encouragement, and criticism over the years. I also would like to thank the NCSM Chem team, specially Dr. Jeff Hammond for quick and clear answering of questions and invaluable help. A CONAC T fellowship (216080) that enabled me to pursue graduate studies for the

degree of Doctor in Chemistry is gratefully acknowledged. I also wish to thank the Instituto de Física and the Dirección General de Tecnologías de la Información y Computación (DITIC) of Universidad Nacional Autónoma de México for excellent and free computer services, in particular to José Luis Gordillo, for his patience and thorough explanations about supercomputing issues.

- ¹I. Shavitt, *J. Mol. Phys.*, **1** (1998).
- ²C. F. Bunge, *J. Chem. Phys.*, **124**, 014107 (2006).
- ³C. F. Bunge and R. Carbó-Dorca, *J. Chem. Phys.*, **125**, 014108 (2006).
- ⁴C. F. Bunge, *J. Mol. Phys.*, **108**, 3279 (2010).
- ⁵T. Shiozaki, Y. Kamiya, S. Hirata, and E. F. Valeev, *J. Chem. Phys.*, **130**, 054101 (2009).
- ⁶S. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and K. A. De Jong, N-Chain: A comprehensive and scalable open-source solution for large scale molecular simulations, *Comput. Phys. Commun.*, **181**, 1477 (2010).
- ⁷J. Tennyson and K. P. Shine, *Philos. Trans. R. Soc. A*, **368**, 2491 (2012).
- ⁸J. K.-S. Chan and J. Head-Gordon, *J. Chem. Phys.*, **119**, 8551 (2003).
- ⁹J. Olsen, P. Jørgensen, H. Koch, A. Balkova, and R. J. Bartlett, *J. Chem. Phys.*, **105**, 8007 (1996).
- ¹⁰J. Bytautas and K. Ruedenberg, *J. Chem. Phys.*, **125**, 174304 (2006).
- ¹¹J. Klopper, *J. Mol. Phys.*, **87**, 481 (2001).
- ¹²J. Noga and J. Klopper, *J. Chem. Phys.*, **115**, 2022 (2001).
- ¹³T. Helgaker, J. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.*, **107**, 9639 (1997).
- ¹⁴P.-O. Löwdin, *Adv. Chem. Phys.*, **6**, 207 (1959).
- ¹⁵C. F. Bunge, *Theor. Chim. Acta*, **1**, 126 (1970).
- ¹⁶T. H. Dunning, Jr., K. A. Peterson, and D. E. Conroy, in *Encyclopedia of Computational Chemistry*, edited by P. von R. Schleyer (Wiley, New York, 1998).
- ¹⁷J. Klopper, K. J. Bak, P. Jørgensen, J. Olsen, and T. Helgaker, *J. Phys. B*, **32**, R103 (1999).
- ¹⁸J. A. Petersson, D. K. Malick, M. J. Frisch, and J. Braunstein, *J. Chem. Phys.*, **122**, 074111 (2005).
- ¹⁹C. F. Bunge, *Phys. Rev.*, **178**, 92 (1968).
- ²⁰A. Bunge, *J. Chem. Phys.*, **42**, 20 (1970).
- ²¹E. P.igner, *Fortran* (Academic, New York, 1959) (English translation).
- ²²A. V. Bunge, C. F. Bunge, R. Jørgensen, and J. Cisneros, *Comput. Chem. Phys.*, **9**, 201 (1989).
- ²³R. Jørgensen, C. F. Bunge, A. V. Bunge, and J. Cisneros, *Comput. Chem. Phys.*, **9**, 223 (1989).
- ²⁴A. V. Bunge, C. F. Bunge, R. Jørgensen, and J. Cisneros, *Comput. Chem. Phys.*, **9**, 239 (1989).
- ²⁵J. Cisneros, R. Jørgensen, C. F. Bunge, and A. V. Bunge, *Comput. Chem. Phys.*, **9**, 255 (1989).
- ²⁶I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, New York, 2009).
- ²⁷R. J. Bartlett and J. Noga, *Rev. Mod. Phys.*, **59**, 291 (2007).
- ²⁸R. E. Brown, Ph.D. thesis, Department of Chemistry, Indiana University, 1967.
- ²⁹C. F. Bunge, *Theor. Chem. Acc.*, **111**, 139 (2010).
- ³⁰A. Halkier, T. Helgaker, P. Jørgensen, J. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *J. Chem. Phys.*, **109**, 243 (1998).
- ³¹A. Chow, J. B. Anderson, and D. Feller, *J. Chem. Phys.*, **109**, 7706 (1997).
- ³²J. R. Alvarez-Idaboy and A. Alano, *Theor. Chem. Acc.*, **111**, 75 (2010).