

APLICACIÓN DE UN ALGORITMO GENÉTICO AL CÁLCULO DE LA FUNCIÓN DE DISTRIBUCIÓN DEL GAS DE ELECTRONES

Vericat Fernando

Grupo de Aplicaciones Matemáticas y Estadísticas de la Facultad de Ingeniería (GAMEFI)
Calle 115 y 49 (1900) La Plata – vericat@iflysib.unlp.edu.ar

Stoico, César O.

Area Física, Facultad de Ciencias Bioquímicas y Farmacéuticas,
Universidad Nacional de Rosario

Carlevaro, C. Manuel

Instituto de Física de Líquidos y Sistemas Biológicos (IFLYSIB)- CONICET La Plata

Palabras claves: Métodos de optimización, algoritmo genético, gas de electrones, función de distribución, aproximación de Jastrow.

Resumen

En este trabajo consideramos la aplicación de un algoritmo genético, como herramienta de optimización, al cálculo de la función de distribución de pares del gas de electrones para la cual proponemos una generalización de la aproximación hipercadena cuántica. La expresión propuesta contiene ciertos parámetros que se determinan minimizando, mediante el algoritmo genético mencionado, la energía del sistema. Comparamos las funciones así obtenidas con las correspondientes calculadas mediante simulaciones Monte Carlo realizadas por otros autores tanto en la versión variacional como difusional. La comparación muestra un muy buen acuerdo especialmente con los resultados de Monte Carlo difusional.

INTRODUCCIÓN

En los últimos años ha habido creciente interés en algoritmos de optimización basados en el principio de evolución de Darwin¹⁻⁷. Algoritmos genéticos, programación evolucionaria, estrategias de juegos, y otras técnicas relacionadas han encontrado un amplio campo de aplicación en ciencia, tecnología, economía, sociología, etc. En particular, los algoritmos genéticos permiten atacar problemas complejos con sorpresiva eficiencia y robustez.

En general, un algoritmo evolucionario está basado en tres preceptos principales:

- i) Es un proceso que se desarrolla a nivel “cromosómico”. Cada individuo se codifica como un conjunto de cromosomas.
- ii) El proceso evoluciona darwinianamente: sobreviven y se reproducen los que mejor se adaptan en un ambiente cambiante.
- iii) El proceso evolucionario tiene lugar en la etapa de reproducción. Es en esta etapa cuando ocurren las *mutaciones* y *entrecruzamientos*. Como resultado los cromosomas de los descendientes pueden diferir de los de los progenitores.

Básicamente, un algoritmo evolucionario construye sucesivas generaciones de individuos, partiendo de una población inicial presupuesta. Cada una de esas generaciones está formada por un conjunto de cromosomas es decir cadenas de símbolos o caracteres (genes) que representan una posible solución al problema bajo consideración. En cada etapa del algoritmo se aplica al conjunto de cromosomas de la correspondiente generación una función de adaptación con el objeto de verificar la bondad de la solución codificada. Entonces se eligen parejas de cromosomas, de acuerdo con su capacidad de adaptación, y se realiza entrecruzamiento o intercambio de algunos de los caracteres que los forman. También, en cada etapa se aplica un operador mutación a algunos de los cromosomas elegidos aleatoriamente.

Tanto para elegir parejas para aplicarles el entrecruzamiento o para elegir los individuos que sufrirán mutación es necesario seleccionar aleatoriamente cromosomas. Los dos métodos más utilizados para realizar esto son:

- i) *Algoritmo de la ruleta*. Consiste en fabricar una ruleta de manera que a cada cromosoma le corresponde un sector proporcional a su capacidad de adaptación.
- ii) *Método del torneo*. Después de mezclar la población se pone a competir sus cromosomas entre ellos en grupos de un dado tamaño (generalmente pares). Los ganadores son los cromosomas con mayor capacidad de adaptación.

Después de esta selección, se procede con la reproducción sexual o cruzamiento de los individuos elegidos. En esta etapa las parejas elegidas intercambian material cromosómico y los cromosomas resultantes pasan a formar parte de la próxima generación. Las formas de reproducción sexual más comúnmente utilizadas son:

- i) *Con un punto de entrecruzamiento*. Este punto es el elegido al azar a lo largo de la cadena y toda la porción de la cadena entre este punto y el final de la misma es intercambiada.
- ii) *Con dos puntos de entrecruzamiento*. La porción a ser intercambiada es aquella comprendida entre los dos puntos de entrecruzamiento que son elegidos al azar.

En la implementación del algoritmo normalmente se asigna al entrecruzamiento un porcentaje que determina la frecuencia de ocurrencia del mismo. Esto significa que no todos los cromosomas intercambian material sino que algunos de ellos pasan intactos a la nueva generación. Al respecto podemos comentar que existe un procedimiento, llamado elitismo, según el cual los individuos que aparecen como mejores adaptados a lo largo de varias generaciones no se cruzan con otros sino que se mantienen intactos hasta que aparece otro individuo mejor adaptado.

Además de la selección y el entrecruzamiento, existe otra operación, la *mutación*, que produce en un cromosoma elegido al azar un cambio en alguno de sus caracteres. Esta operación permite introducir nuevo material cromosómico en la población. Al igual que para el entrecruzamiento, la mutación se maneja con un porcentaje que determina la frecuencia con que ocurrirá. Este porcentaje es generalmente menor al 5%, bastante menor que el correspondiente porcentaje para el entrecruzamiento.

Una vez que los cromosomas elegidos han sido entrecruzados y mutados, necesitamos de algún método de sustitución. Esto es, entre todos los individuos de la población debemos elegir aquéllos que serán sustituidos por la nueva progenie. Existen dos métodos de sustitución que son los más utilizados. En uno de ellos todos los padres son sustituidos por los nuevos individuos generados. De esta manera un individuo nunca coexiste con sus padres. El otro solamente los individuos peor adaptados de la población son sustituidos, permitiendo de esta manera la coexistencia entre padres y progenie.

Puesto que obviamente casi siempre desconocemos la respuesta a nuestro problema, debemos establecer un criterio para detener el algoritmo. Podemos mencionar dos de tales criterios: i) el algoritmo se corre hasta alcanzar un número dado de generaciones; ii) el algoritmo se termina cuando se logra equilibrar la población, esto es, cuando todos o al menos la mayoría de los individuos tiene la misma capacidad de adaptación.

En este trabajo aplicamos alguna de las ideas anteriores con la finalidad de calcular la función de distribución de pares (PDF) para el gas de electrones homogéneo.

El modelo de gas de electrones homogéneo de Sommerfeld⁸ ha demostrado ser de gran utilidad para describir muchas de las propiedades electrónicas de los sólidos metálicos. El modelo representa a los electrones de conducción como un conjunto de fermiones puntuales cargados a temperatura cero moviéndose en un continuo neutralizante que juega el papel de la

red iónica. En la versión más simple del modelo los fermiones son considerados sin spin y el continuo neutralizante es caracterizado simplemente por una constante dieléctrica. Si consideramos N fermiones de masa m cada uno con una carga e , entonces el Hamiltoniano del sistema es

$$H = \left(\frac{\hbar^2}{2m} \right) \sum_{i=1}^N \nabla_i^2 + \sum_{i<j} v(r_{ij}) \quad (1)$$

En esta expresión $v(r_{ij})$ es el potencial de pares dado por

$$v(r_{ij}) = \frac{e^2}{\epsilon r},$$

(2) donde $r_{ij} = |\vec{r}_i - \vec{r}_j|$, con \vec{r}_i indicamos la posición de la i -ésima partícula y con ϵ la constante dieléctrica que en este trabajo supondremos que vale 1.

El objeto central en nuestra teoría son las funciones de distribución de pares (PDF)^{9,10}. Estas funciones establecen un nexo entre el comportamiento microscópico y la termodinámica del sistema. Si denotamos con $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ a la función de onda de N partículas, entonces la función de distribución de pares viene dada por

$$\begin{aligned} \rho(\vec{r}_1, \vec{r}_2) &= \rho(\vec{r}_1) \rho(\vec{r}_2) g(\vec{r}_1, \vec{r}_2) \\ &= N(N-1) \frac{\int \dots \int d\vec{r}_3 d\vec{r}_4 \dots d\vec{r}_N |\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2}{\int \dots \int d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_N |\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2} \end{aligned} \quad (3)$$

donde la integración es sobre todo el volumen y

$$\rho(\vec{r}_1) = N \frac{\int \dots \int d\vec{r}_2 d\vec{r}_3 \dots d\vec{r}_N |\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2}{\int \dots \int d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_N |\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2}$$

(4) La función $g(\vec{r}_1, \vec{r}_2)$ es la función de correlación de pares (PCF). Para sistemas homogéneos (como consideraremos aquí), $\rho(\vec{r}_i)$ es la densidad numérica de electrones $\rho(\vec{r}_i) = \rho = N/V$ (V = volumen del sistema) y entonces tanto la PDF como la PCF dependen sólo de la separación de las partículas: $\rho(r_{12}) = \rho^2 g(r_{12})$.

Muchas de las teorías del gas homogéneo de electrones se apoyan en el principio variacional según el cual la energía del estado fundamental E_0 es una cota mínima para el valor medio del Hamiltoniano calculado usando cualquier función de onda aproximada de prueba $\psi_T(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$:

$$\frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \geq E_0 \quad (5)$$

Frecuentemente se utiliza como función de onda aproximada una forma factorizada

$$\psi_T(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = F(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \psi_0(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \quad (6)$$

En esta expresión $\psi_0(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ es la función de onda del sistema cuando las interacciones son anuladas ($v(r_{ij}) \equiv 0$). Se trata de una función asimétrica frente a las permutaciones de partículas y como tal generalmente se utiliza un determinante de Slater. El factor simétrico

$F(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ da cuenta de las correlaciones entre las partículas cuando las interacciones son conectadas. De acuerdo con Jastrow¹¹, el factor de correlación de N partículas puede ser factorizado:

$$F(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \prod_{i < j} f(\vec{r}_i, \vec{r}_j) \quad (7)$$

PDF parametrizada

Supondremos que la función de onda de prueba para el sistema de N electrones sin spin cuyo Hamiltoniano está dado por Ecs.(1-2) tiene la forma de Ec.(6) donde para la parte ideal $\psi_0(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ usamos una generalización parametrizada de una expresión debida a Lado¹² y para los factores de correlación de Jastrow usamos también una expresión parametrizada que contiene un pseudo potencial calculado en la aproximación de fases al azar (RPA)^{13,14}. Concretamente proponemos

$$\psi_T(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \exp\left\{\frac{1}{2} \sum_{i < j} [\alpha w_0(r_{ij}) - \beta u_{RPA}(r_{ij})]\right\}$$

(8) donde α y β son los parámetros a ajustar.

En la Ec. (8), el potencial efectivo para el gas ideal, $w_0(r)$, viene dado por¹²

$$w_0(r) = \ln g_0(r) - \frac{1}{(2\pi)^3 \rho} \int d\vec{k} e^{-\vec{k} \cdot \vec{r}} \frac{[S_0(k) - 1]^2}{S_0(k)}$$

(9) En esta expresión $g_0(r)$ y $S_0(k)$ son la PCF y el factor de estructura para el gas ideal, respectivamente, cuyas expresiones son^{15,16}:

$$g_0(r) = 1 - \frac{9}{2} \left[\frac{\text{sen}(k_F r) - k_F r \cos(k_F r)}{(k_F r)^3} \right]^2$$

(10) y

$$S_0(k) = \begin{cases} 1 & k > 2k_F \\ \frac{3}{4} \frac{k}{k_F} - \frac{1}{16} \left(\frac{k}{k_F} \right)^3 & k < 2k_F \end{cases} \quad (11)$$

donde k_F es el momento de Fermi: $k_F = (3\pi^2 \rho^*)^{\frac{1}{3}}$ con $\rho^* = \rho a_0$ ($a_0 =$ radio de Bohr).

El pseudopotencial RPA, por otra parte, se escribe en el espacio de momentos:

$$2\rho^* \tilde{u}_{RPA}(k) = -\frac{1}{S_0(k)} + \left[\frac{1}{S_0(k)^2} + \frac{4m\rho^* \tilde{v}(k)}{\hbar^2 k^2} \right]$$

(12) donde, en general, $\tilde{f}(k)$ indica transformada de Fourier de $f(r)$.

En principio para calcular la PDF con Ec.(3) debemos integrar el cuadrado de la función de onda sobre $3N$ coordenadas. Se trata obviamente de un problema realmente formidable

desde el punto de vista práctico dado que se supone N muy grande (en el límite termodinámico $N \rightarrow \infty$). Para obviar este inconveniente hacemos uso de una forma modificada de la aproximación hipercadena (HNC) para la cual la PDF se escribe como una ecuación integral simple involucrando solamente un par de partículas. Ignorando los llamados diagramas elementales (que son muy difíciles de calcular) escribimos¹⁷

$$g(r) = \exp[\alpha w_0(r) - \beta u_{RPA}(r) + N(r)] \quad (13)$$

$$N(r) = \rho \int [g(r') - 1 - N(r')] [g(|\vec{r} - \vec{r}'|) - 1] d\vec{r}',$$

(14) donde $N(r)$ indica la suma de los diagramas nodales.

Para hacer uso de la aproximación variacional implícita en la Ec.(5) necesitamos una expresión para el valor medio del Hamiltoniano. Haciendo uso de la identidad de Jackson-Feenberg¹⁸ obtenemos

$$\frac{E}{N} = \rho \int d\vec{r} g(r) \left[-\frac{\hbar^2}{4m} \nabla^2 \ln f^2(r) + v(r) \right]. \quad (15)$$

Optimización de los parámetros

Nuestro problema es resolver la ecuación integral dada por las Ecs.(13 y 14) con los parámetros α y β determinados de manera que minimicen la energía escrita como funcional de $g(r)$ en la Ec. (15). Con este fin implementamos un algoritmo genético.

Primeramente procedemos a generar la población inicial, la cual está formada por N_p réplicas aleatorias de la cadena de dos números (la cual representa un individuo o cromosoma de la población) $\gamma^{(\alpha)}, \gamma^{(\beta)}$ donde $\gamma^{(\alpha)}$ es un número real elegido al azar en el intervalo $[0,1]$ y redondeado a un número n de decimales establecido, el cual es asignado al parámetro α . Dada la cadena $\gamma^{(\alpha)}, \gamma^{(\beta)}$, la *codificación* consiste en reemplazar la secuencia de números reales por un solo número natural obtenido poniendo las partes decimales una a continuación de la otra. Luego si $\gamma^{(\alpha)} = 0.\gamma_1^{(\alpha)}\gamma_2^{(\alpha)} \dots \gamma_n^{(\alpha)}$ y $\gamma^{(\beta)} = 0.\gamma_1^{(\beta)}\gamma_2^{(\beta)} \dots \gamma_n^{(\beta)}$ entonces obtenemos la cadena $\gamma_1^{(\alpha)}\gamma_2^{(\alpha)} \dots \gamma_n^{(\alpha)}\gamma_1^{(\beta)}\gamma_2^{(\beta)} \dots \gamma_n^{(\beta)}$ y la población es el conjunto

$$\left\{ \left(\gamma_1^{(\alpha)}\gamma_2^{(\alpha)} \dots \gamma_n^{(\alpha)}\gamma_1^{(\beta)}\gamma_2^{(\beta)} \dots \gamma_n^{(\beta)} \right)_r \quad r = 1, 2, \dots, N_p \right\}$$

En términos genéticos la codificación produce la estructura cromosómica de los individuos. El proceso inverso a la codificación es la *decodificación* que devuelve los parámetros α y β correspondientes a cada individuo. En la decodificación permitimos que los parámetros devueltos sean multiplicados por un factor constante $\eta > 1$ con el objeto que éstos puedan tomar valores en un intervalo más amplio que $[0,1]$.

Como función de adaptación del r -ésimo individuo usamos $f_r = e^{-E_r}$ donde E_r es la energía calculada con la Ec.(15) cuando $g(r)$ es calculada de resolver la ecuación integral definida por

las Ecs.(13 y 14) para los parámetros α y β obtenidos de decodificar la estructura cromosómica del r -ésimo individuo de las población.

El cálculo continúa dividiendo la población de N_p réplicas en $N_p/2$ parejas que hemos elegido aleatoriamente mediante el algoritmo de la ruleta².

Una vez que la primera generación de réplicas (padres) ha sido generada y dividida en parejas, la segunda generación puede ser generada aplicando entrecruzamientos entre los miembros de cada pareja. De vez en cuando algunos de los miembros de la nueva generación de réplicas pueden ser cambiados aplicando el operador mutación.

Dada una pareja, el entrecruzamiento se efectúa generando un nuevo número aleatorio $c \in [0,1]$ el cual es comparado con una probabilidad de entrecruzamiento pre-establecida $p \in [0,1]$. Si $c \leq p$ el operador entrecruzamiento actúa intercambiando todos los dígitos desde la s -ésima posición hasta el final de la cadena entre los miembros de la pareja. Aquí s es un entero al azar tal que $1 \leq s \leq 2n$. Por ejemplo, si la pareja está formada por las cadenas 153280472...26337 y 768325399...85069 y es $s=4$ entonces la nueva pareja será 153280472...25069 y 768325399...86337.

Para aplicar el operador mutación primero elegimos al azar aquellas cadenas que mutarán. Entonces, por cada una de ellas se elige aleatoriamente un dígito (gen) y se lo cambia por otro número natural $\ell \in [0,9]$ elegido al azar.

RESULTADOS

Como puede fácilmente apreciarse, el gas de electrones queda completamente determinado dando solamente su densidad ρ o, como es costumbre en la teoría de muchos-cuerpos, dando el radio de Wigner-Seitz $r_s = [3/4\pi\rho]^{1/3}$. En este trabajo hemos aplicado el procedimiento descrito anteriormente a un gas de electrones homogéneo a densidades metálicas $1 \leq r_s \leq 10$ y también a $r_s = 50$. En nuestros cálculos usamos $N_p = 150$ y $n = 5$.

En la Tabla I mostramos los parámetros α y β obtenidos para los diversos valores de r_s considerados. Debemos notar el cambio en la tendencia para los valores más grandes de r_s . Justamente para valores mayores que $r_s = 50$ comenzamos a tener severos problemas de convergencia.

Tabla I: parámetros α y β

r_s	α	β
1	2.22	1.64
2	7.87	14.86
3	9.67	30.52
6	14.71	43.03
10	44.97	18.46
50	-6834.17	128.71

Las figuras 1-3 muestran las correspondientes $g(r)$'s. En la Fig. 1 hemos puesto juntas todas las curvas que hemos obtenido. Observamos las características salientes del gas de electrones: cuando la densidad aumenta (r_s disminuye) el comportamiento tiende a la de un

gas de fermiones paramagnético ideal con valor al contacto $\frac{1}{2}$ y alcanzando rápidamente su valor asintótico 1. Cuando la densidad disminuye las funciones de correlación se vuelven más estructuradas mostrando, en particular, un agujero de Pauli más pronunciado cerca del origen.

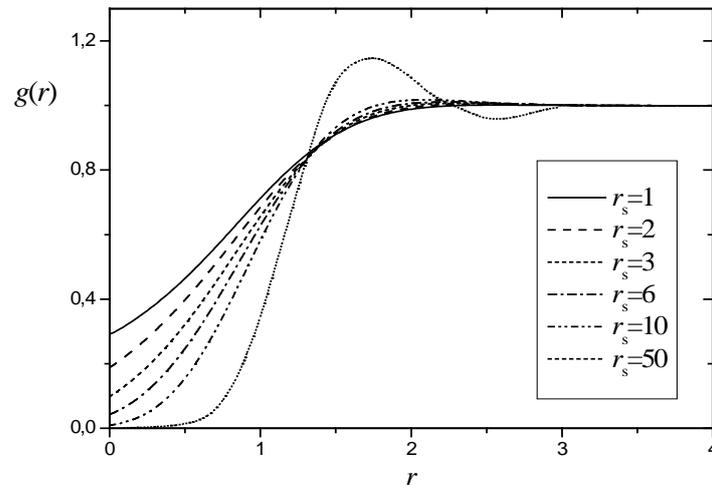


Fig. 1 Las funciones $g(r)$ del gas de electrones homogéneos para $r_s=1,2,3,6,10$ y 50 obtenidas en este trabajo.

En la Figs. 2 y 3 comparamos nuestros resultados con los obtenidos por otros autores mediante simulaciones Monte Carlo en las variantes difusional (DMC) y variacional (VMC)¹⁹⁻²¹.

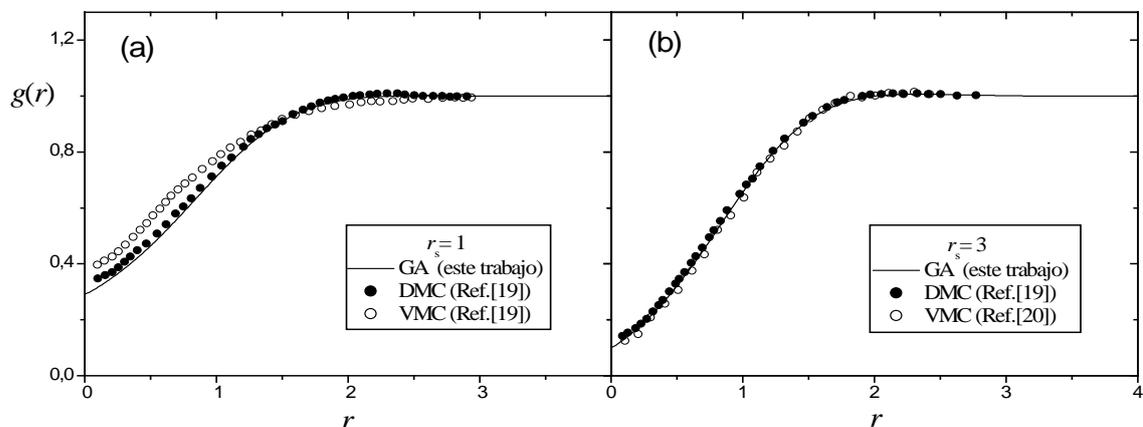


Fig. 2 Comparación de las funciones $g(r)$ del gas de electrones obtenidas en este trabajo con las calculadas mediante Monte Carlo difusional y variacional por otros autores. (a) $r_s=1$; (b) $r_s=3$

Un primer aspecto a resaltar es la diferencia que existe entre los resultados de Monte Carlo entre sus versiones difusional y variacional especialmente para valores pequeños de r_s . También debemos remarcar el buen acuerdo entre los resultados de este trabajo y las simulaciones Monte Carlo, particularmente DMC. Merece señalarse que el tiempo para obtener una de nuestras curvas haciendo correr el algoritmo completo con una Pentium IV es del orden de 2 a 3 días para las densidades metálicas. Estos tiempos son comparables con las

simulaciones DMC que fueron realizadas según sus autores con una supercomputadora CRAY-YMP¹⁹.

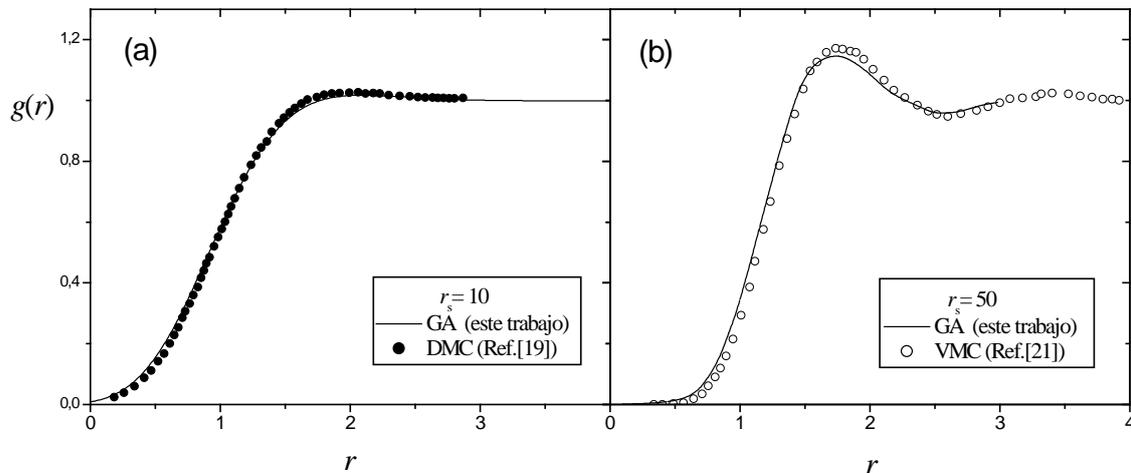


Fig. 3 Comparación de las funciones $g(r)$ del gas de electrones obtenidas en este trabajo con las calculadas mediante Monte Carlo difusional y variacional por otros autores. (a) $r_s=10$; (b) $r_s=50$

AGRADECIMIENTOS

Este trabajo fue parcialmente subvencionado por Universidad Nacional de La Plata (Proyecto 11/I153), Universidad Nacional de Rosario (Proyecto 19/VET47), Consejo Nacional de Investigaciones Científicas y Técnicas (PIP 1192) y Agencia Nacional de Promoción Científica y Tecnológica (PICT 00908). F.V. y C.M.C. son miembros del CONICET.

Referencias

1. J.H. Holland, *Adaptation in Natural and Artificial Systems* (University of Michigan, Ann Arbor, MI, 1975).
2. D.E. Goldberg, *Genetic Algorithm in Search, Optimization and Machine Learning* (Addison-Wesley, 3. Reading, MA, 1989).
4. M. Mitchell, *An Introduction to Genetic Algorithm* (Prentice Hall, 1998).
5. D. Fogel, *Evolutionary Computation* (IEEE Press, 1996).
6. T. Bäck, *Evolutionary Algorithm in Theory and Practice* (Oxford Press, 1996).
7. L. Davis (Ed.), *Genetic Algorithms and Simulated Annealing* (Pitman. London. 1987).
8. L. Hoddeson, G. Baym and M. Eckert, *Rev. Mod. Phys.* 59, 287 (1987).
9. G.D. Mahan, *Many-Particle Physics* (Plenum, New York, 1981).
10. N.H. March and M.P. Tosi, *Coulomb Liquids* (Academic, London, 1984).
11. R.J. Jastrow, *Phys. Rev.*, 98, 1479 (1955).
12. F. Lado, *J. Chem. Phys.* 47, 5369 (1967).
13. D.M. Ceperley, *Phys. Rev. B* 18, 3126 (1978).
14. T. Gaskell, *Proc. Phys. Soc.* 77, 1182 (1961); 80, 1091 (1962).
15. G. Placzek, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, California, 1951).
16. F. London, *J. Chem. Phys.*, 11, 203 (1943).
17. C.O. Stoico, C.M. Carlevaro, D.G. Renzi and F. Vericat, *Physica E*, 42, 1691 (2010).
18. Feenberg, *Theory of Quantum Fluids* (Academic Press, New York, 1969).
19. G. Ortiz and P. Ballone, *Phys. Rev. B*, 50, 1391 (1994); erratum: *Phys. Rev. B.*, 56, 9970 (1997).
20. L.J. Lantto, *Phys. Rev. B*, 22, 1380 (1980).
21. D.M. Ceperley and B.J. Alder, *Journal de Physique*, 41, C7-295 (1980).