

固定节面量子 Monte Carlo 的一个新算法

黄宏新

曹泽星

(湖南师范大学化学系,长沙 410081)

(厦门大学化学系,厦门 361005)

摘要 提出了一个固定节面量子 Monte Carlo 的新算法,与前人的算法相比,其“指导函数”的优化不是在扩散前,而是在扩散过程中同步进行;这不仅在机时上是节省的,更重要的是优化与扩散两者按相同的取样方式进行,达到互相改善的目的;这一优化方案是超线性收敛的,它能加快粒子的扩散;在扩散过程中,“指导函数”的节面不断发生改变,这有利于减小“固定节面误差”。这一新算法已被运用到 CH_2 的 X^3B_1 态和 a^1A_1 态,以及 NH_2 的 $\pi-X^2B_1$ 态和 $\sigma-A^2A_1$ 态总能量的计算,由此算出了 CH_2 单一三重态的“劈开”能 $\Delta E_{S-\tau}=(45.542 \pm 1.840)\text{kJ/mol}$ 和 NH_2 的 $\sigma-\pi$ “劈开”能 $\Delta E_{\sigma-\pi}=(141.644 \pm 1.589)\text{kJ/mol}$. 计算结果表明这一新算法在精度、统计误差和计算量方面比一般固定节面量子 Monte Carlo 方法都要优越得多。

关键词 固定节面量子 Monte Carlo 方法 扩散方程 指导函数 超线性收敛

在各种求解 Schrödinger 方程的 Monte Carlo 方案中,固定节面量子 Monte Carlo (FN-QMC) 方法是最常用的一种,关于 FNQMC 的评述请参考文献[1,2]. 但 FNQMC 方法存在如下缺点:(1)在开始 FNQMC 计算前,其“指导函数”中的大量参数必须先通过变分 Monte Carlo (VMC) 方法优化^[3],优化这些参数是很费机时的;而且效果并不佳,因为 FNQMC 和 VMC 的取样方式根本不同.(2)前人在优化量子 Monte Carlo (QMC) 波函数方面有如下工作:Umrigar 在一套固定的构型中优化波函数^[4];实践表明,这种方案在波函数节面附近会出现“权重失控”而使优化无法进行下去^[5];Huang^[3]和 Bueckert^[5]均采用传统“速降法”实施优化,但传统速降法优化速度慢,还常常因为出现病态而找不到最优点^[6].

最近我们提出了一个优化 QMC 波函数的新方法^[7],这是一个步长自动调节、超线性收敛的方法. 不过这个方法仍然是在 VMC 过程中实施的,因此对 FNQMC 计算并不很合适. 本文将对这一新的优化方法进行改造,使之适合于 FNQMC 的取样方式;将导出扩散条件下的优化原理,建立自优化-扩散过程,将扩散与优化有机结合,形成一个新的 FNQMC 算法. 这个算法有如下优点:(1)它是一个优化、扩散“一体化”的过程,在这个过程中“指导函数”的优化与粒子的扩散同步进行. 这不仅在机时上是节省的,更重要的是优化与扩散两者按相同的取样方式进行,达到互相改善的目的,因而这是一个自改善的 Monte Carlo 方案.(2)与一般的 FN-QMC 方法不同,在我们的自优化-扩散过程中,“指导函数”的节面不断发生改变,这有利于减小“固定节面误差”. (3)因为这一优化方法本身是一个具有超线性收敛速度的方法,因此能

1996-04-03 收稿,1996-07-23 收修改稿

加快 FNQMC 的扩散过程。(4)在优化过程中,我们使用分析导数,而不是用差分代替,因此没有经验参数。

为了检验这一新算法的性能,我们用它处理了 CH_2 的单一三重态能级“劈开”(the singlet-triplet splitting in CH_2)和 NH_2 的 σ - π “劈开”(the σ - π splitting in NH_2);并将处理结果与其他量子化学方法(如全 CI)的处理结果作了对照。检验结果表明,我们提出的这种自优化-扩散量子 Monte Carlo (SODMC)方法是非常成功的。实际上,对于像 CH_2 和 NH_2 这样的分子,由于它们的“小体积”,还由于他们的化学活性取决于其基态与激发态的“靠近度”,因此对这些基团的“劈开”能的研究一直是各种计算方法进行比较的很好算例。

1 FNQMC 方法^[1,2]

引入 $f(R, t) = \Phi(R, t) \Psi_T(R)$, 这里 Φ 是体系的真实波函数, Ψ_T 是重要性取样的试探函数(又称指导函数), f 是几率分布函数;含虚时间的 t 的 Schrödinger 方程的扩散方程形式为:

$$\frac{\partial}{\partial t} f(R, t) = \frac{1}{2} \nabla^2 f(R, t) - \nabla \left[\frac{1}{\Psi_T(R)} \nabla \Psi_T(R) \right] f(R, t) + [E_T - E_L(R)] f(R, t), \quad (1)$$

这里 $E_L = H\Psi_T/\Psi_T$ 称为局部能; E_T 是一个试探能量。扩散子由时间 t 到 $t + \tau$ (τ 是一个微时间差)的扩散由下列 Green 函数决定:

$$G(R \rightarrow R', \tau) = (2\pi\tau)^{-3n/2} \exp\{-\tau[E_L(R') - E_T]\} \times \exp\left(\frac{-[R - R' - \tau \nabla \ln |\Psi_T(R)|]^2}{2\tau}\right), \quad (2)$$

于是在 R' 处的几率分布函数为:

$$f(R', t + \tau) = \int G(R \rightarrow R', \tau) f(R, t) dR. \quad (3)$$

反复对扩散子用(2)、(3)式进行扩散,直至达到稳定态。由稳定态的构型即可计算各种力学量。

2 扩散条件下的优化原理

在 FNQMC 计算过程中,粒子的漂移是在试探函数 Ψ_T 的指导下进行的, Ψ_T 的质量好坏直接关系到 FNQMC 的精度、收敛速度和统计误差的大小。我们发现对指导函数 Ψ_T 进行优化最有效的办法是极小化 Δ :

$$\Delta = \frac{1}{N} \sum_{i=1}^N [E_L - E_T]_i^2 = \overline{(E_L - E_T)^2}. \quad (4)$$

这里 N 是扩散过程中的构型数。

记 $\Psi_T = \Psi_T(P, R)$, P 是 J 个待优化参数的列向量, $P = (P_1, P_2, \dots, P_J)$ 。 Δ 对 P 的一阶、二阶偏导数为:

$$g = \frac{\partial \Delta}{\partial P} = \frac{1}{N} \sum_{i=1}^N [2(E_L - E_T) \frac{\partial E_L}{\partial P}]_i, = \overline{2(E_L - E_T) \frac{\partial E_L}{\partial P}}. \quad (5)$$

$$q = \frac{\partial^2 \Delta}{\partial P^2} = \frac{1}{N} \sum_{i=1}^N 2 \left[\left(\frac{\partial E_L}{\partial P} \right)^2 + (E_L - E_T) \frac{\partial^2 E_L}{\partial P^2} \right]_i \approx \frac{1}{N} \sum_{i=1}^N 2 \left[\left(\frac{\partial E_L}{\partial P} \right)^2 \right]_i = 2 \overline{\left(\frac{\partial E_L}{\partial P} \right)^2} \geq 0. \quad (6)$$

(6)式中取近似的理由见文献[7], 注意这里 g 和 q 均为列向量.

我们在以前的文章中^[7]指出了极小化 $\Delta(P)$ 的下降方案为:

$$P_{k+1} = P_k - \lambda_k \left(\frac{g}{q} \right)_k, \quad (7)$$

这里 k 表示第 k 次重复搜索, λ_k 是步长因子, 它由下式自动确定:

$$\lambda_k = \frac{(P_k - P_{k-1})^T \cdot (g_k - g_{k-1})}{(g_k - g_{k-1})^2}. \quad (8)$$

(8)式中的 $(P_k - P_{k-1})^T$ 表示列向量 $(P_k - P_{k-1})$ 的转置. 我们已经证明^[7], 这一下降方案是一个没有经验参数的, 收敛速度极快的方法. 它比前人的优化方法快 3~5 倍, 且精度和稳定性也要好得多.

下面我们将这个下降方案与扩散过程结合起来, 由(5)~(8)式, 有:

$$P_{k+1} = P_k - \frac{1}{2} \frac{(P_k - P_{k-1})^T \cdot \left\{ [(E_L - E_T) \frac{\partial E_L}{\partial P}]_k - [(E_L - E_T) \frac{\partial E_L}{\partial P}]_{k-1} \right\}}{\left\{ [(E_L - E_T) \frac{\partial E_L}{\partial P}]_k - [(E_L - E_T) \frac{\partial E_L}{\partial P}]_{k-1} \right\}^2} \times \left\{ \frac{(E_L - E_T) \frac{\partial E_L}{\partial P}}{\left(\frac{\partial E_L}{\partial P} \right)^2} \right\}_k. \quad (9)$$

(9)式即为扩散条件下优化波函数的基本公式.

3 自优化-扩散过程

在开始 FNQMC 计算前, 必须选择好试探函数 Ψ_T 的形式、参数 P 和试探能 E_T 的初值. 这方面的内容请参考文献[4, 7]. 自优化-扩散过程为:

(A) 由 $\Psi_T(P_{k=0}, R)$ 选取 N ($N=3000 \sim 5000$) 个扩散子的初始构型, 我们称它为构型块 0; 在构型块 0 中用(5)、(6)式求出 $(g/q)_{k=0}$ 的值; 令 $\lambda_0=0.05 \sim 0.1$, 更新参数 P 的值:

$$P_{k=1} = P_{k=0} - \lambda_0 \left(\frac{g}{q} \right)_{k=0}. \quad (10)$$

由此得到新的 $\Psi_T(P_{k=1}, R)$.

(B) 从构型块中取出第 m 个构型, 这个构型里的电子将运动; 它们在微时间 τ 内按 Green 函数(2)式中的 Gauss 部分独立地扩散. 如果移动的是第 m 个构型中的第 j 个电子, 它将运动到:

$$R_j^{(m)} = R_j^{(m)} + \tau \nabla \ln |\Psi_T(R)| + \chi, \quad (11)$$

这里 $R_j^{(m)}$ 是电子 j 的三维坐标矢量, χ 是一个均值为 0、方差为 τ 的三维 Gauss 随机变量. 当电子 j 按(11)式移动后, 接受这个移动的概率为:

$$A(R \rightarrow R', \tau) = \min\left(1, \frac{|\Psi_T(R')|^2 G(R' \rightarrow R, \tau)}{|\Psi_T(R)|^2 G(R \rightarrow R', \tau)}\right), \quad (12)$$

当第 m 个构型里的 n 个电子都移动 1 次后, 计算构型 m 的分枝率 M_m , 它由(2)式中的指数因子决定:

$$M_m = \exp\{-\tau[E_L(R') - E_T]\}. \quad (13)$$

当构型块中的 N 个构型都按(11)~(13)式扩散后, 一个新的构型块(块 1)已经生成.

(C) 注意到(9)式等号右边的量均为构型块 0 和构型块 1 所决定, 因此对块 0 和块 1 使用(9)式即可求出 $P_{k=2}$ 的值, 由此得到一个更新的 $\Psi_T(P_{k=2}, R)$.

(D) 对体系反复使用步骤(B)、(C), 直至达到稳定态.

在计算过程中注意以下几点:(1)试探能 E_T 的值由前 1 个块中 E_L 的平均值担任, 所以 E_T 的更新与块的更新同步;(2)在循环过程中, 若出现 $\lambda \leq 0$, 则令 $\lambda = 0.05 \sim 0.1$, 再进入循环过程.(3)我们使用 $|\Delta E_T| \leq 10^{-5}$ 为收敛准则.

4 结果和讨论

我们用这一新的 FNQMC 算法计算了 CH_2 的 X^3B_1 态和 a^1A_1 态的总能量, 由此算出了两者的能量差, 即单一三重态的“劈开”能 ΔE_{S-T} , 计算结果列于表 1. 对于 NH_2 的 $\pi-X^2B_1$ 态和 $\sigma-A^2A_1$ 态也作了相应处理, 求出了 $\Delta E_{\sigma-\pi}$, 计算结果列于表 2, 计算中我们取 $N = 4\,000$, $\Delta t = 0.000\,1$ (a. u.), 试探函数形式见文献[7].

表 1 不同方法获得的 CH_2 的 X^3B_1 态和 a^1A_1 态的相关能及单一三重态“劈开”能 ΔE_{S-T} 数据比较^{a)}

	X^3B_1	a^1A_1	$\Delta E_{S-T}/\text{kJ}\cdot\text{mol}^{-1}$
	E 相关(a. u.)	E 相关(a. u.)	
SCF	0	0	109.192
FCI	0.118 313	0.140 886	50.059
BEN/CASNO	0.119 920	0.147 392	37.178
BMP/SONO	0.112 853	0.129 801	64.779
一般 FNQMC	0.111 ± 0.001	0.132 ± 0.003	53.697 ± 7.862
SODMC	$0.121\,1 \pm 0.000\,2$	$0.145\,4 \pm 0.000\,7$	45.542 ± 1.840

a) 几何构型: X^3B_1 : $\text{CH} = 0.108\,2$ nm, $(\text{HCH}) = 132.4^\circ$; a^1A_1 : $\text{CH} = 0.111\,6$ nm, $(\text{HCH}) = 102.4^\circ$

表 2 不同方法获得的 NH_2 的 $\pi-X^2B_1$ 态和 $\sigma-A^2A_1$ 态的相关能及 $\sigma-\pi$ “劈开”能 $\Delta E_{\sigma-\pi}$ 数据比较^{a)}

	$\pi-X^2B_1$	$\sigma-A^2A_1$	$\Delta E_{\sigma-\pi}/\text{kJ}\cdot\text{mol}^{-1}$
	E 相关(a. u.)	E 相关(a. u.)	
SCF	0	0	130.646
FCI	0.169 621	0.165 570	141.226
BEN/CASNO	0.184 447	0.185 776	127.175
BMP/SONO	0.163 037	0.160 063	138.424
一般 FNQMC	0.164 ± 0.002	0.159 ± 0.004	143.735 ± 10.497
SODMC	$0.170\,1 \pm 0.000\,4$	$0.165\,9 \pm 0.000\,6$	141.644 ± 1.589

a) 几何构型: $\pi-X^2B_1$: $\text{NH} = 0.102\,4$ nm, $(\text{HNH}) = 103.4^\circ$; $\sigma-A^2A_1$: $\text{NH} = 0.100\,0$ nm, $(\text{HNH}) = 144.0^\circ$

为了便于比较,我们还对这 2 个分子进行了一般 FNQMC 计算(分子构型和试探函数形式均与 SODMC 相同),计算结果列于各自的表中. 表中还列出了前人用 SCF, FCI, BEN 和 BMP 方法处理的相应的数据,这些数据来源于文献[8].

由表 1 的数据可知,我们的算法(SODMC)对于 X^3B_1 和 a^1A_1 态而言分别获得了 0.121 1 和 0.145 4 (a. u.) 的相关能;统计误差很小,仅为 0.000 2 和 0.000 7 (a. u.). 而一般 FNQMC 方法则获得 0.111 和 0.132 (a. u.) 的相关能;统计误差为 0.001 和 0.003 (a. u.). 对于单一三重态“劈开”能 ΔE_{S-T} , SCF, FCI, BEN 和 BMP 的值分别为 109.192, 50.059, 37.178 和 64.779 kJ/mol. SODMC 的值为 (45.542 ± 1.840) kJ/mol, 与全 CI 的值接近;而一般 FNQMC 的值为 (53.697 ± 7.862) kJ/mol, 统计误差过大,不能说明问题. 对于表 2 中的数据,也可作类似分析.

总之, SODMC 方法与一般的 FNQMC 相比,具有精度高、统计误差小的特点,另外,计算量也比一般 FNQMC 小. 例如对于 CH_2 的单一三重态“劈开”能 ΔE_{S-T} 的计算,一台 486 微机, SODMC 方法和一般 FNQMC 方法的 CPU 时间分别为 38 和 59 min, 而 NH_2 的 $\Delta E_{\sigma-\pi}$ 的计算,其 CPU 时间分别为 47 和 71 min;这说明本算法是非常成功的.

参 考 文 献

- 1 Ceperley D, Alder B. Quantum Monte Carlo. *Science*, 1986, 231: 555~560
- 2 Anderson J B. Fixed-node quantum Monte Carlo. *International Reviews in Physical Chemistry*, 1995, 14 (1): 85~112
- 3 Huang S, Sun Z. Optimized trial function for quantum monte carlo. *J Chem Phys*, 1990, 92: 597~602
- 4 Umrigar C J, Wilson K G, Wilkins J W. Optimized trial wavefunction for quantum Monte Carlo calculation. *Phys Rev Lett*, 1988, 60: 1 719~1 722
- 5 Bueckert H, Rothstein S M, Vrbik J. Optimization of quantum Monte Carlo wavefunction using analytical derivatives. *Can J Chem*, 1992, 70: 366~371
- 6 Akaike H. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Ann Inst Statist Math Tokyo*, 1959, 11: 1~17
- 7 Huang Hongxin, Cao Zexing. A novel method for optimizing quantum Monte Carlo wavefunctions. *J Chem Phys*, 1996, 104: 200~205
- 8 Parisel O, Ellinger Y. Second-order perturbation theory using correlated orbitals. *Chemical Physics*, 1994, 189: 1~16