

# 用降低固定样点组上的能量方差改进量子 Monte Carlo 方法中的试探波函数\*

孙 祉 伟

黄 升 培

(中国科学院力学研究所, 北京 100080) (加利福尼亚大学化学系, 贝克来)

王 德 民

(北京大学技术物理系, 北京 100871)

## 摘 要

本文在量子 Monte Carlo 方法中用固定样点组降低能量方差对试探波函数的参数进行改进, 并将以此方法优化后的试探波函数, 用于  $H_2$ ,  $Li_2$  和  $H_2O$  分子能量计算, 获得很好的结果。

关键词: 量子 Monte Carlo, 固定样点, 试探波函数

## 一、引 言

作者在文献[1]中已指出: 量子 Monte Carlo 方法 (QMC) 的显著优点是它可以高精度地计算体系能量, 一般能得到 90% 到 100% 的相关能, 而且, 它的计算程序非常简单, 内存要求少, 又不受波函数形式的限制。然而, 在量子 Monte Carlo 方法中试探波函数  $\psi_T$  的作用极为重要<sup>[2]</sup>。一方面, 它直接决定了计算的准确度; 另一面, 它影响到收敛速度。通常选取试探波函数的办法是直接取自洽场近似的结果  $\phi(R)$  乘上电子-电子相关函数及电子-核相关函数, 即

$$\psi_T = \phi(R)S_c S_n$$

这样的波函数是应该重新优化的。我们曾在随机行走过程中找出变分能量或其方差的变化趋势, 用最陡下降法在不断的“行走”中得到波函数参数的优化值。“随机行走”在很大程度上避免了取样的影响, 但它又使得“目标函数”的估算带有随机误差, 即不能在有限计算时间中给出确定的值。

Bartlett<sup>[3]</sup> 早在 1935 年就用降低一组固定样点局域能方差的办法去检验和改进波函数的质量。Coldwell 在 1977 年也用过类似的方法改进波函数。Umrigar 等很成功地将类似方法用于单原子分子和离子的 Monte Carlo 变分能量的计算<sup>[4,5]</sup>。采用固定样点改进波函数的突出优点是: 它只用位形空间 (Configuration Space) 中为数不多的取样点来检验波函

本文 1989 年 12 月 4 日收到, 1990 年 4 月 6 日收到修改稿。

\* 国家自然科学基金资助项目。

数的品质。显然,这会大大减少计算量,不必计算变分能量的积分,此时“目标函数”(能量偏差)的数值是完全确定的,无随机误差,因而在优化过程中可以运用更有效的方法(如牛顿法一类具有二级收敛速率,比只有一级速率的最陡下降法有效得多)。

虽然用固定样点优化波函数看来十分诱人,但由于样点十分有限,人们自然很关心优化结果究竟如何受取样点影响?例如,样点数目如何选取以及参照能的选择如何影响结果等。本文将讨论这些问题,并把该法推广到二原子和三原子分子,取得了很好的结果。一般,改进波函数包括两方面:一是改进波函数的函数形式本身;二是优化函数中的参数。本文的重点是优化参数,为了更有效地表明优化的机制和规律,我们的计算实例只用了较简单的函数形式。

## 二、优化方法

### 1. 方法的表述

当波函数  $\psi_T$  接近能量算符 Hamilton 的一个本征函数  $\phi_i$  时,局域能应在整个位形空间接近  $\phi_i$  的本征值。这一事实使人们想到作为一个好的波函数,对于任何一组固定样点,其局域能的方差应很小。因而人们可以用缩小这一方差的办法来改进波函数。方差为

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (E_L(R_i) - E_r)^2, \quad (1)$$

其中  $E_r$  为参照能,它与所计算的能本征值很接近;求和是对  $N$  个取样点作的;  $R_i$  为第  $i$  个样点的坐标,即体系中  $n$  个电子位置坐标的总和。这些样点的分布是根据  $|\phi_0|^2$  来取样的( $\phi_0$  是起始试探波函数)。因为在优化的过程中波函数  $\psi_T$  是在不断变化的,最好根据  $\psi_T$  (即时试探波函数)来调节点的分布。如果仍用  $|\phi_0|^2$  的分布点,则可用下式表示:

$$\sigma^2 = \frac{1}{N} \frac{\sum_{i=1}^N \left(\frac{\psi_T}{\phi_0}\right)^2 (E(R_i) - E_r)^2}{\sum_{i=1}^N \left(\frac{\psi_T}{\phi_0}\right)^2}. \quad (2)$$

这是 Umrigar 所建议的式子,看起来似乎更合理。然而,在实用中可发现,当用收敛速度较快的优化方法时,此式会使优化过程不稳定,效果不好。这是因为  $\psi_T$  变化太大时,使得  $\frac{\psi_T}{\phi_0}$  与 1 相差很大,由  $\psi_T$  到  $\phi_0$  的变换已失去实际意义。所以本文采用(1)式。

### 2. 取样点的数目和选法

显然,样点取得越多,越有利于正确确定波函数。但从提高效率来看,选样点越少越好。最理想的情况是刚好选了精确的函数形式,如果该函数有  $p$  个参数,那末只要选  $p$  个样点就可以完全确定这  $p$  个参数而得到正确的本征函数,并且与点的选取方式无关。当然,一般这是不可能的。可以想到,取样点的数目与所研究的体系大小有关。电子数越多,越能组成更多的分布特征,因而需要更多的样点才能代表这些特征。另外,与体系的对称性也有关,对称性高的体系需要较少的样点;对称性低的体系则需要较多样点。例如,球对称的单原子分子,当整个电子体系绕原子核转动时并不构成新的分布特征。 $H_2$  分子有轴对称,用 200—500 个样点就够了;  $Li_2$  分子至少需要 1000 个样点;而有 10 个电子的  $H_2O$  分子则需要 2000 个样点(40 个待优化变分参数),与此对比,Umrigar 等发现,同样具 10 个电子的 Ne 原子,有 50 个参数,

却只要 1000 个样点。其重要原因是 Ne 原子具球对称性。当然,样点的数目,原则上也应与选用的函数形式有关,函数形式越正确,需要的样点数目越少。

在选样点上,应当考虑如何把有限个样点尽量能代表出电子的分布特征。分布密度大的地方应多取样点。这些样点的分布是按  $|\phi_0|^2$  选定的,为此就要进行以  $\phi_0$  为重要性函数的变分 Monte Carlo 计算。最后,还要对样点进行检查,去掉那些局域能与平均局域能相差很大的样点,因为这些具有大偏离的局域能的样点会把优化过程导入歧途,造成不良结果。这些极少数样点比例很小,其影响却很大。因为总样点数目有限,只能把最有代表性的样点选入,而去掉那些不太“合格”的代表。

当完成基于分布  $|\phi_0|^2$  样点的优化步骤后,可以得到波函数  $\phi_1$ , 然后再以  $|\phi_1|^2$  来选样点,重复优化过程,得到进一步改进的波函数  $\phi_2$ 。当然,还可以一直做下去。一般,到  $\phi_1$  就可以了。

### 3. 参照能的选取

首先,要确定优化波函数的判据是什么。这里所说的波函数的优劣是以其相应的变分能量的低高来衡量的。也就是变分能量越低的波函数品质越好。那么能否通过固定样点去降低变分能量的途径以改进波函数质量呢? 这是不行的。因为样点数目有限,总可以调整参数使在样点上的局域能很低,甚至还可以比真实能量还低。由于真正的变分能量是要通过整个空间积分而求得的,如果对固定样点强行降低能量,则在其它点(非样点)的能量必然升高,以致真正的变分能量并不低。如果对固定样点降低局域能的方差,则情况与上述的不同。它是使每个样点的局域能都趋于一个参照能值,而不是无限降低能值,因而可以真正改善波函数的品质。由此可见,参照能  $E_r$  的选取对优化结果是有直接影响的。

假设  $E_r = E + \delta$ , 这里  $E$  是所考虑的态的本征值。于是有

$$\sigma^2 = \frac{1}{N} \sum_i (E_L - E_r)^2 = \sigma_0^2 - 2\delta \langle E_L \rangle + \delta^2 + 2\delta E, \quad (3)$$

其中  $\sigma_0^2$  是局域能与本征能的方差,即

$$\sigma_0^2 = \frac{1}{N} \sum_i (E_L - E)^2,$$

$\langle E_L \rangle$  是平均局域能,即

$$\langle E_L \rangle = \frac{1}{N} \sum_i E_L(i).$$

(3)式的第三、四两项是常量。若选  $E_r < E$ , 则  $\delta < 0$ , 欲减少  $\sigma^2$ , 可以通过降低局域能与本征能方差  $\sigma_0$  和  $\langle E_L \rangle$  来实现; 若选  $E_r > E$ , 则  $\delta > 0$ , 欲减少  $\sigma^2$ , 可以通过降低  $\sigma_0^2$  及增加  $\langle E_L \rangle$  来实现。根据经验,选  $E_r$  比本征能略高为好。因为若选  $E_r < E$ ,  $\delta < 0$ , 当  $(E_L - E)^2$  不能进一步减少时,  $\langle E_L \rangle$  项将使结果变坏。对于  $H_2$  分子,  $E_r$  的选取对优化结果影响不太大。但对  $Li_2$  和  $H_2O$  分子,  $E_r$  的选取就很重要了。事实上, (3) 式第二项的影响取决于  $\delta$  和  $\langle E_L \rangle$ 。对于  $H_2$ ,  $E \approx -1.174$  a. u., 无论选  $E_r = -1.18$  a. u. 或  $-1.16$  a. u., 其  $\delta = \pm 0.01$ , 它们在第二项中的值都很小, 所以选  $E_r > E$  或  $E_r < E$  都可以, 影响不大。然而, 对于  $Li_2$  和  $H_2O$  分子, 它们的  $\langle E_L \rangle$  分别约  $-14$  a. u. 和  $-76$  a. u., 这对第二项的影响就很大了:

#### 4. 局域极小值

至今所使用的绝大多数优化方法只能得到局域极小值。为了得到较好的结果,在优化过程中应注意两点:(1)局域极小值强烈地依赖参数的起始值,所以有时故意略为变动这些参数,看能否改进原来的结果;(2)因为优化过程受多种因素影响(如  $E_r$  和样点的选取等),可以在不同条件下做优化,看能否找到更好的结果。

#### 5. 迭代收敛

多参数优化的各种方法都是一种迭代过程。因此,在开始优化之前,先需设置一个  $\sigma^2$  的容许值(即到达该值后就不再迭代了)或最大迭代次数。作为一般的规律,  $\sigma^2$  在前 20—30 次迭代中下降明显,之后就减慢了。收敛到很小的  $\sigma^2$ ,并不总意味着波函数品质更好(即变分能更低),因为这与样点和  $E_r$  的选取有关。如样点和  $E_r$  的选取不合适,更小的  $\sigma^2$  也会导致更差的结果。

### 三、试探波函数的形式和优化参数的选取

本文选用了比较简单的试探波函数形式,它由分子轨道的 Slater 行列式乘以电子-电子核和电子-核的相关函数,即

$$\psi_T = |D_{i1}^{\alpha}| |D_{i1}^{\beta}| S_e S_n, \quad (4)$$

其中  $|D_{i1}^{\alpha}|$  是由电子自旋为  $i$  的分子轨道组成的 Slater 行列式。  $S_e$  和  $S_n$  分别是电子-电子及电子-核相关函数。对于  $S_e$ ,我们采用下列形式<sup>[6]</sup>:

$$S_e = \exp \left[ -b \sum_{ij} \exp(-a\gamma_{ij}) \right]. \quad (5)$$

为了满足尖顶条件 (Cusp Condition)  $\left. \frac{\partial \psi}{\partial \gamma_{ij}} \right|_{\gamma_{ij}=0} = \psi/2$  (对自旋相反的电子对),应取  $a = 0.5/b$ 。这样,上述相关函数只包含一个变分参数。这个函数形式很便于运算(尤其是微商运算),并且相当好地描述了电子的相关性<sup>[6]</sup>。

对于  $S_n$ ,应满足的尖顶条件是  $\left. \frac{\partial \psi}{\partial \gamma_{i\alpha}} \right|_{\gamma_{i\alpha}=0} = -Z\psi$ 。同时,为了增加试探波函数的可调性,本文采用通常的 Jastrow 形式,即

$$S_n = \exp \left[ \sum_{i,\alpha} \frac{-\lambda \gamma_{i\alpha}}{1 - \nu \gamma_{i\alpha}} \right], \quad (6)$$

其中  $\gamma_{i\alpha}$  是电子  $i$  与核  $\alpha$  的距离。

分子轨道由 Slater 型原子轨道线性组合而成。下面以  $\text{Li}_2$  分子为例说明波函数的参数考虑。 $\text{Li}_2$  分子的 3 个分子轨道由 STO 原子轨道线性组合如下:

$$\begin{aligned} \phi_k = & C_{k1}\phi_{1s}(1) + C_{k2}\phi_{2s}(1) + C_{k3}\phi_{2p_x}(1) \\ & + C_{k4}\phi_{1s}(2) + C_{k5}\phi_{2s}(2) + C_{k6}\phi_{2p_x}(2). \end{aligned} \quad (7)$$

考虑到  $\text{Li}_2$  的对称性,对于态  $\text{Li}_2(X^1\Sigma_g^+)$ ,各系数间有下列关系:

$$\begin{aligned} C_{14} = C_{11}, \quad C_{15} = C_{12}, \quad C_{16} = -C_{13}, \\ C_{24} = -C_{21}, \quad C_{25} = -C_{22}, \quad C_{26} = C_{23}, \\ C_{34} = C_{31}, \quad C_{35} = C_{32}, \quad C_{36} = -C_{33}, \end{aligned} \quad (8)$$

其中系数  $C_{11}$ ,  $C_{21}$  和  $C_{31}$  可固定,其余 6 个系数待优化;还有 6 个原子轨道  $\phi_{1s}$ ,  $\phi_{2s}$  和  $\phi_{2p_x}$  的

指数 ( $\phi_1$  与  $\phi_2$  取相同值;  $\phi_3$  取其它值) 以及相关函数中的  $b, \lambda$  和  $\nu$ , 共 15 个参数。

对  $H_2$  分子(态  $X^1\Sigma_g^+$ ), 有 2 个系数; 3 个轨道指数以及  $b, \lambda, \nu$ , 共 8 个参数。  $H_2O$  分子共用了 40 个参数: 15 个系数, 20 个指数和 5 个相关参数。

### 四、优化过程和计算结果

由于样点是固定的, 作为优化过程的“目标函数”的  $\sigma^2$  有着完全确定的值, 所以可采用具有二级收敛速率的优化方法, 如 Newton 或准 Newton 法。又由于(1)式是平方和形式, 故采用非线性最小二乘法效率更高。为了给出较精确的微分估算值, 最好用双精度计算, 并用中心差分技术算一阶导数效果较好。图 1—4 给出了  $H_2$  分子在 43 次迭代过程中各物理量和参数的变化。为了对比它们在迭代过程中的变化趋势, 纵坐标标度不便统一标出, 而将它们各自的最大值、最小值和近 40 次以后迭代的平稳值列于表 1 (本文用了 500 个样点)。

表 1 能量、方差和各参数在迭代过程的极值和平稳值

	$\langle E \rangle$ (a. u.)	$\sigma^2$	$C_{12}$	$C_{13}$	$\xi_1$ (1s)	$\xi_2$ (2s)	$\xi_3$ (2p <sub>r</sub> )	$b$	$\lambda$	$\nu$
最大值	-1.15520	0.045943	0.0300	0.0770	1.4469	1.7576	2.0000	1.0451	0.2000	1.0592
最小值	-1.17286	0.007514	-0.5331	0.0037	0.9087	1.5953	0.4964	0.5000	-0.0914	0.0611
平稳值	-1.16771	0.007514	0.019	0.0390	0.9087	1.5953	0.4964	1.0139	0.2696	0.0611

从表 1 中可以看出, 方差  $\sigma^2$  一直在减少, 平均能量  $\langle E \rangle$  最低曾到 -1.17286 a. u., 但到最

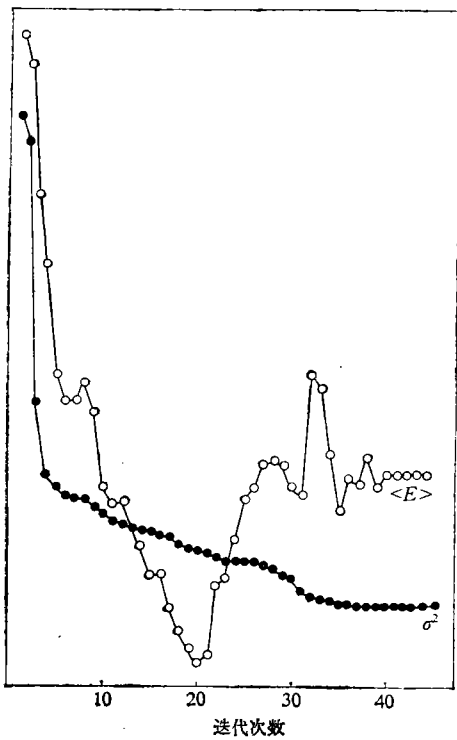


图 1  $\sigma^2$  和  $\langle E \rangle$  随迭代次数的变化

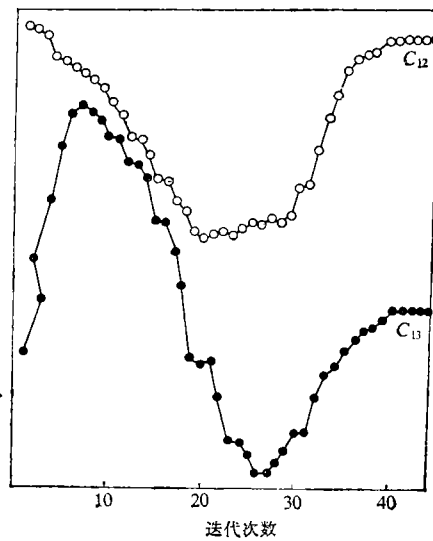
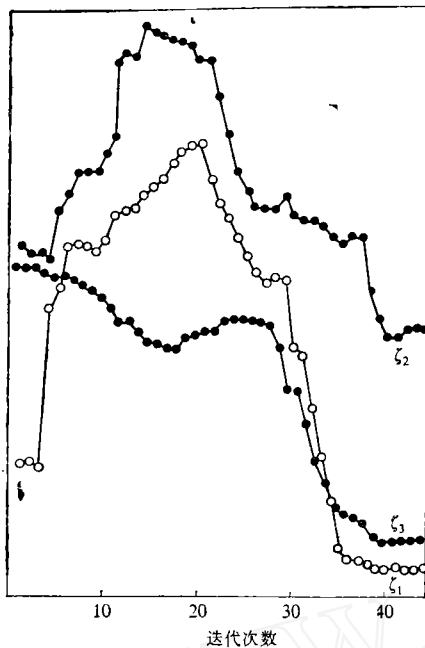
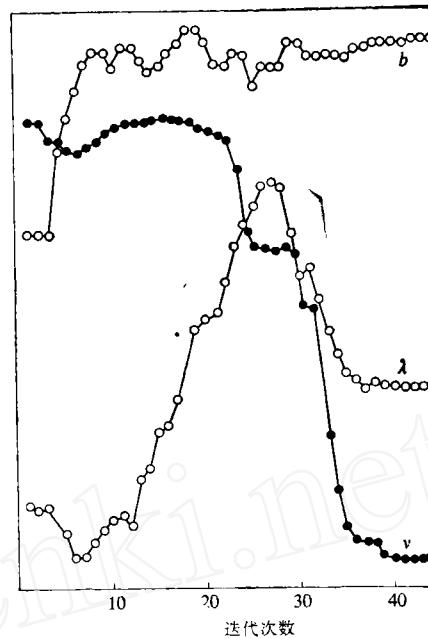


图 2  $C_{12}$  和  $C_{13}$  随迭代次数的变化

图 3  $\zeta_1$ ,  $\zeta_2$  和  $\zeta_3$  随迭代次数的变化图 4  $b$ ,  $\lambda$  和  $\nu$  随迭代次数的变化

后又有所上升,用此组参数单独算的变分能量为  $-1.172$  a. u.

我们用最好的优化结果给出  $H_2$  分子变分能量为  $-1.1724 (\pm 0.0001)$  a. u., 亦即获得了约 95% 相关能 (HF 能量为  $-1.1336$ ). 因为  $H_2$  分子基态没有节面, 所以固定节面 QMC 能量应能得到  $H_2$  基态的精确能值 ( $-1.17447$  a. u.). 本文实际得到的是  $-1.1745$  a. u.

对于  $Li_2$  分子的计算, 变分能值为  $-14.956$  a. u. 即获得 69% 相关能 (HF 能量为  $-14.872$  a. u., 精确能量为  $-14.995$  a. u.). QMC 能值为  $-14.992$ , 即得到约 98% 的相关能.

对于  $H_2O$  分子, 其 HF 和精确能值分别为  $-76.0675$  a. u. 和  $-76.4376$  a. u.. 本文用优化后的波函数计算得到变分能和固定节面能分别为  $-76.238$  a. u. 和  $-76.403$  a. u., 即分别获得 47% 和 91% 的相关能.

由于优化的目标是降低能量的方差, 也即在优化过程中总是在调节参数使得各点的局域能趋于正确值, 这样就可以推想, 这一过程也有利于尖顶条件的满足 (也即消除  $\frac{1}{r}$  的奇点). 我们把优化后得到的波函数做了这种检验, 发现它们确实能自动地在很大程度上满足尖顶条件. 尖顶条件的满足无疑会对加快计算的收敛速度起到极重要的作用, 因而也是一般衡量波函数好坏的标准之一.

## 五、结 论

在计算中, 我们用了比较简单的函数形式, 经优化后, 对  $H_2$ ,  $Li_2$  和  $H_2O$  来说, 无论是变分能或 QMC 能都有很大改进. 实际上是至今所发表的最好结果. 这就表明用固定样点优化

波函数是行之有效的方法。与 Umrigar 相比,他只算了原子的变分能量,而本文则算了二原子和三原子分子的变分能量及固定节面能量。

固定样点优化法的优点是明显的。但人们自然关心优化结果是否依赖于取样?从本文的工作来看,对取样的依赖性存在的。有的较明显,有的较隐蔽,只有注意控制这些因素,才能得到最好结果。

本文所讨论的波函数优化法只限于优化波函数的参数。为了全面改进波函数,还应在波函数的形式上有所改进。例如,本文用单参数的相关函数形式和并不复杂的基底函数,用变分 Monte Carlo 法就可得到  $H_2$  分子相关能的 95%;可是对  $Li_2$  分子只能得到 69% 相关能;对  $H_2O$  则只能得到 47%。这说明,随着分子的原子数目的增加,相关函数显然应考虑更多的因素。现用的单参数的形式只与两电子的距离有关,这是一种最粗略的描述。文献[6]中已证明电子对的位置也起作用,特别是当电子数增加时就需要更精确地描述电子的相关性。

由于 QMC (指固定节点 Monte Carlo 或 Green function Monte Carlo) 计算通常要比变分 MC 多费 1—2 个量级的机时(对达到相同精度而言)。要使后者的精度能与前者相比,必须在改进波函数方面下大功夫。

#### 参 考 文 献

- [1] 孙祉伟,王德民,化学物理学报, **2**(1989), 243.
- [2] Reynolds, P. J. et al., *J. Chem. Phys.*, **77**(1982), 5593.
- [3] Bartlett, J. H. et al., *Phys. Rev.*, **47**(1935), 679.
- [4] Umrigar, C. J. et al., *Phys. Rev. Lett.*, **60**(1988), 1719
- [5] Landau, D. P. & Schlutler, H. B., *Computer Simulation in Condensed Matter Physics: Recent Developments*, Springer-Verlag, 1988.
- [6] Sun Zhiwei, et al., *Theor. Chim. Acta*, **75**(1989), 353.