

# 推广模拟退火方法及其应用

向 阳 龚新高

(中国科学院固体物理研究所,合肥 230031)

**摘 要:** 本文介绍了推广模拟退火方法的基本思想及其统计基础。我们通过一系列标准函数测试了推广模拟退火方法的性能。作为该方法的应用,我们研究了 Thomson 问题和 Ni 团簇结构。最后,探讨了推广模拟退火方法的效率随体系复杂性的变化。

**关键词:** 最优化,推广模拟退火算法,Thomson 问题

**中图分类号:** O224 ;O411.9 **文献标识码:** A

## 1 引 言

在自然科学,社会科学,以及人们的日常生活中广泛存在着大量的求最大或最小的问题,即所谓的最优化问题。用数学语言来说,就是决定一组参量,使其对应的目标函数达到最小值或最大值。例如,近年来,由于  $C_{60}$  的发现,物理学家和化学家对团簇发生了广泛而浓厚的兴趣<sup>[1~9]</sup>。真实团簇的结构,就是指其能量(自由能)最低的结构,所以求团簇结构问题,也就是求其能量最低的结构问题。近年来自旋玻璃的研究中也涉及到求体系总能最低的结构问题。事实上,最优化问题遍及理论物理和实验物理的各个领域<sup>[10]</sup>,有必要对最优化问题和最优化方法进行广泛而深入的研究。

总体上最优化方法可以分为两大类:决定性方法(determined methods)和随机方法(stochastic methods)<sup>[13]</sup>。决定性的方法可以非常快,但容易陷入局部极小(local minimum)。相反,随机方法不易陷入局部极小,但在有限步数内没有一种随机方法可以保证找到全局极小。在所有的优化方法中,模拟退火方法被认为是最有效的方法之一。

模拟退火方法的基本思想来源于固体的退火过程。首先我们回顾一下固体退火过程的物理图象<sup>[12]</sup>。在加热固体时,固体中原子的热运动不断增强,随着温度的不断升高,固体的长程有序被彻底破坏,固体熔解为液体(或气体)。冷却时,液体中原子的热运动渐渐减弱,随着温度的徐徐降低,原子运动渐趋有序。当温度降至结晶温度后,原子运动变为围绕晶体格点的微小振动,液体凝固成固体,这种由高温向低温逐渐降温的过程称为退

火。退火过程中系统的熵值不断减小,系统能量随温度降低趋于最小值。

对固体退火过程的研究给人们以新的启示。1982年,Kirkpatrick等人<sup>[11]</sup>首先意识到固体退火过程与优化问题之间存在的类似性,Metropolis等对固体在恒定温度下达到热平衡过程的模拟也给他们以启迪:应该把Metropolis准则引入到优化过程中来。最终他们得到一种对Metropolis算法进行迭代的优化算法,这种算法类似固体退火过程,称之为“模拟退火算法”。具体方案实施如下:

(1) 初始时刻  $t = 1$ , 在足够高的初始温度  $T(t)$  下, 随机选取一个初始态  $x_t$ , 计算  $E(X_t)$

(2) 根据某种跃迁分布, 随机地从  $X_t$  跳到  $X_{t+1}$

(3) 计算  $E(X_{t+1})$ , 如果  $E(X_{t+1}) < E(x_t)$ , 则用  $X_{t+1}$  取代  $X_t$ , 否则计算接收几率  $p = \exp\{-[E(X_{t+1}) - E(X_t)]/T(t)\}$ , 且产生一个  $(0, 1)$  内均匀分布的随机数  $r$ , 比较  $p$  与  $r$ , 如果  $p > r$ , 则用  $X_{t+1}$  取代  $X_t$ , 否则保持不变。

(4) 根据降温公式, 计算下一步的温度  $T(t+1)$ , 回到(2)重复以上过程, 直到  $E$  达到某个预定的精度范围。

在以上方案实施过程中, 在任意一下温度下, 体系可以重复试探  $L$  次, 这里  $L$  称为 MapKoB 链长。

根据跃迁分布, 降温方式, 接收几率的不同, 模拟退火可以分为三种形式: 经典模拟退火(CSA), 快速模拟退火(FSA), 推广模拟退火(GSA)。

在经典模拟退火(CSA)<sup>[11]</sup>, 跃迁分布是高斯形式的:

$$g(\Delta x) \propto \exp\left(-\frac{(\Delta x)^2}{T}\right) \quad (1)$$

这里  $\Delta x$  是变量  $x$  的试探跃迁距离,  $T$  是约化单位的退火温度。如果跳跃使新能量变低, 则新态接收, 否则就用 Boltzman-Gibbs 形的接收几率来判断接收。接收几率用 Metropolis 形式:

$$p = \min\left[1, \exp\left(-\frac{\Delta E}{T}\right)\right] \quad (2)$$

CSA 可以收敛到一个不错的全局极小, 但正如我们以下要指出的, 它的收敛相当慢。Geman 等人指出, 如果用以上形式的跃迁和接收几率, 则以几率 1 找到全局极小的充要条件是温度随时间对数下降<sup>[14]</sup>, 这就是所谓的经典模拟退火, 又称为 Boltzmann 机器。

1987年, Szu 和 Hartley 提出了所谓的快速模拟退火(FSA)<sup>[15]</sup>。FSA 使用半局域的 Cauchy-Lorentz 跃迁分布:

$$g(\Delta x) \propto \frac{T}{(T^2 + (\Delta x)^2)^{\frac{D+1}{2}}} \quad (3)$$

这里  $D$  是变量空间的维数。CSA 采用的高斯分布是一个局域分布, 而 FSA 采用的 Cauchy-Lorentz 跃迁分布是一个半局域的分布, 在同样的温度下, Cauchy-Lorentz 分布使体系有更大机会进行长距离跃迁。FSA 中的温度随时间的倒数  $1/t$ , 其接收几率仍用 Metropolis 的形式。

以上介绍了 CSA 和 FSA。尽管在目标函数的变量较少, 局域极小不太多时它们能给

出合理的结果,当目标函数的变量很多,局域极小也很多时,这些方法容易陷入局部极小而找不到全局极小<sup>[13]</sup>,为此人们又发展了更有效的推广模拟退火算法<sup>[16~18]</sup>。

2 推广模拟退火的统计基础-Tsallis 统计

Tsallis 将熵的表达式  $s = -\kappa \sum_{i=1}^W p_i \ln p_i$  (此处  $W$  是可能态的总数,  $p_i$  是相应的几率)推广为<sup>[19]</sup>:

$$s_q = \kappa \frac{1 - \sum_{i=1}^W p_i^q}{q - 1}, \quad (q \in \mathbb{R}) \tag{4}$$

可看出,当  $q \rightarrow 1$  时,  $s_q$  恢复原来熵的表达式。利用  $\sum_{i=1}^W p_i = 1$ ,  $s_q$  也可以写成:

$$s_q = \frac{\kappa}{q - 1} \sum_{i=1}^W p_i (1 - p_i^{q-1}) \tag{5}$$

(5)式表明,在所有情况下  $s_q \geq 0$ ,当  $W = 1$  时,  $s_q = 0$ 。

在正则系统中,在两个约束  $\sum_{i=1}^W p_i = 1$  和  $\sum_{i=1}^W p_i \epsilon_i = U_q$  ( $\epsilon_i$  是能级)下求极值可得到:

$$p_i = \frac{[1 - \beta(1 - q)\epsilon_i]^{\frac{1}{1-q}}}{Z_q} \tag{6}$$

和

$$Z_p \equiv \sum_i [1 - \beta(1 - q)\epsilon_i]^{\frac{1}{1-q}} \tag{7}$$

$Z_p$  是推广配分函数,  $\beta = 1/\kappa T$  是 Lagrange 参数。在  $q \rightarrow 1$  极限下,

$$p_i = e^{-\beta \epsilon_i} / Z_1 \tag{8}$$

$$Z_1 = \sum_{i=1}^W e^{-\beta \epsilon_i} \tag{9}$$

又恢复了 Boltzmann 统计。

图 1 是几率  $p$  随能量变化图像。1)当  $q < 1$  时,对所有的  $\epsilon \geq \frac{1}{\beta(1-q)}$ ,其几率  $p = 0$ ,即存在一个 cutoff =  $\frac{1}{\beta(1-q)}$ 。在这个 cutoff 处,当  $q = 0$  时,斜率为  $-1$ ,当  $q < 0$  时,斜率为  $-\infty$ ,当  $q > 0$  时,斜率为  $0$ 。2)当  $q = 1$  时,几率分布恢复了 Boltzmann 形式;3)当  $q > 1$  时,在  $\epsilon = \frac{1}{\beta(1-q)}$  处,几率分布发散,当  $\beta \epsilon \rightarrow +\infty$  时,  $p \rightarrow 0$ 。

以上是关于 Tsallis 统计的一些基本知识,包括 Tsallis 熵, Tsallis 几率分布等等。Tsallis 统计,特别是 Tsallis 几率分布,是推广模拟退火的基础。

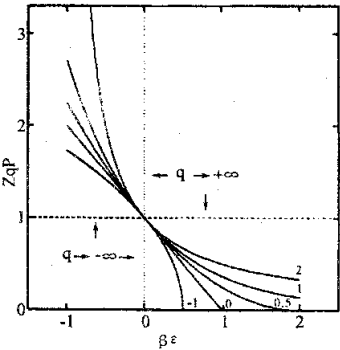


图 1 几率 P 随能量的变化

3 推广模拟退火的提出

在模拟退火<sup>[11,15]</sup>中,最重要的三个部分就是降温方式、跃迁几率分布和接收几率分布。本节的主要内容是给出它们在推广模拟退火<sup>[16~18]</sup>中的表达式<sup>[16]</sup>。

在 Boltzmann 机器中,接收几率用了 Metropolis 方案:

$$P_1(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1}) = \begin{cases} 1 & E(\mathbf{X}_{t+1}) < E(\mathbf{X}_t) \\ e^{[E(\mathbf{X}_t) - E(\mathbf{X}_{t+1})]/T_1^a(t)} & E(\mathbf{X}_{t+1}) \geq E(\mathbf{X}_t) \end{cases} \quad (10)$$

当  $T_1^a(t) = 0$  如果  $E(\mathbf{X}_{t+1}) < E(\mathbf{X}_t)$ , 则  $P_1 = 1$  如果  $E(\mathbf{X}_{t+1}) \geq E(\mathbf{X}_t)$ , 则  $P_1 = 0$ 。以上限制对退火过程的渐进平衡很重要,在推广形式中也应满足。根据推广的 Tsallis 几率,推广的接收准则被写成:

$$P_{q_a}(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1}) = \begin{cases} 1 & E(\mathbf{X}_{t+1}) < E(\mathbf{X}_t) \\ \frac{1}{[1 + (q_a - 1)(E(\mathbf{X}_t) - E(\mathbf{X}_{t+1}))/T_{q_a}^a(t)]^{\frac{1}{q_a-1}}} & E(\mathbf{X}_{t+1}) \geq E(\mathbf{X}_t) \end{cases} \quad (11)$$

为简单起见,我们假设  $E(\mathbf{X}) \geq \alpha \ (\forall \mathbf{X})$ , 而且,我们也假设  $q_a \geq 1$ , 因此  $T_{q_a}^a(t)$  能降到 0 而不产生任何奇异性。在以上假设下,  $P_{q_a} \in [0, 1] \ (\forall q_a)$  且对  $E(\mathbf{X}_{t+1}) \geq E(\mathbf{X}_t)$ , 随着温度  $T_{q_a}^a(t)$  从  $+\infty$  下降到 0,  $P_{q_a}$  能单调地从 1 到 0。当  $E(\mathbf{X}_{t+1}) < E(\mathbf{X}_t)$  时,  $P_{q_a} = 1$ 。当  $q_a < 1$  时, 如果  $[1 + (q_a - 1)(E(\mathbf{X}_{t+1}) - E(\mathbf{X}_t))/T_{q_a}^a(t)] < 0$ , 则令  $P_{q_a} = 0$ 。

现在来给出各向同性的跃迁几率分布的表达式  $g_{q_v}(\Delta x_t)$  是控制跃迁几率分布的一个参量, 和降温方式的表达式, 其中  $\Delta x_t = (\mathbf{X}_{t+1} - \mathbf{X}_t)$ , 设跃迁几率分布为:

$$g_{q_v}(\Delta x_t) = c \frac{[T_{q_v}^v(t)]^d}{\{[T_{q_v}^v(t)]^e + (q_v - 1)(\Delta x_t)\}^{\frac{a}{q_v-1}}} \quad (12)$$

其中  $a, b, c, d, e$  是待定的参数, 它们与  $(q_v, D)$  有关。跃迁几率分布满足归一化条件:

$$\Omega_D \int_0^\infty d\rho \rho^{D-1} g_{q_v}(\rho) = 1 \quad (13)$$

此处  $\Omega_D \equiv D\pi^{D/2}/\Gamma(\frac{D}{2} + 1)$  是  $D$  维空间立体角。当  $q_v = 1$  时, 要求此跃迁几率分布恢复经典模拟退火所用的高斯分布:

$$g_1(\Delta x_t) \propto \exp\left[-\frac{(\Delta x_t)^2}{T_1^v(t)}\right] \quad (14)$$

此处  $T_1^v(t)$  是  $q_v = 1$  的跃迁温度。把方程 (14) 代入到方程 (13), 利用欧拉积分  $\Gamma(\alpha) = 2 \int_0^\infty t^{2\alpha-1} e^{-t^2} dt$  可确定归一化系数为  $[\pi T_1^v(t)]^{-D/2}$ , 则方程 (14) 可写为:

$$g_1(\Delta x_t) = \frac{e^{-(\Delta x_t)^2/T_1^v(t)}}{[\pi T_1^v(t)]^{D/2}} \quad (15)$$

当  $q_v = 2$  时, 要求此跃迁几率分布恢复快速模拟退火所用的 Cauchy-Lorentz 分布:

$$g_{\Sigma}(\Delta x_t) \propto \frac{T_{\Sigma}^v(t)}{\{[T_{\Sigma}^v(t)]^2 + (\Delta x_t)^2\}^{\frac{D+1}{2}}} \quad (16)$$

此处  $T_{\Sigma}^v(t)$  是  $q_v = 2$  的跃迁温度。把方程 (16) 代入到方程 (13), 利用欧拉积分  $B(a, b) = \int_0^{\infty} \frac{x^{a-1}}{(1+x)^{a+b}} dx$ , 以及  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  可确定归一化系数为  $\Gamma(\frac{D+1}{2})/\pi^{\frac{D+1}{2}}$ , 则方程 (16) 可写为:

$$g_{\Sigma}(\Delta x_t) = \frac{\Gamma(\frac{D+1}{2})}{\pi^{\frac{D+1}{2}}} \frac{T_{\Sigma}^v(t)}{\{[T_{\Sigma}^v(t)]^2 + (\Delta x_t)^2\}^{\frac{D+1}{2}}} \quad (17)$$

对方程 (12) 利用归一化方程 (13) 并且考虑  $\Delta x_t$  可能包括多个维数, 可得到

$$d = e^{\frac{2a - D(q_v - 1)}{2(q_v - 1)}} \quad (\forall q_v, \forall D) \quad (18)$$

在  $t \rightarrow \infty$  的极限下, 体系必须收敛到全局极小。当这种情况发生时, 态的跃迁必须是 “infinite often in time (IOT)”<sup>[15]</sup>, 这就要求  $\sum_{t=t_0}^{\infty} g_{q_v}(\Delta x_{t_0})$  在  $t_0 \gg 1$  时对固定的  $\Delta x_{t_0}$  发散。在这些条件下, 有:

$$\sum_{t=t_0}^{\infty} g_{q_v}(\Delta x_{t_0}) \propto \sum_{t=t_0}^{\infty} [T_{q_v}^v(t)]^d \quad (19)$$

我们知道, 在经典模拟退火中温度随时间对数下降, 在快速模拟退火中温度随时间的倒数  $1/t$  变化, 即, 对任意  $D$ ,  $T_1^v(t) = T_1^v(1) \ln 2 / \ln(1+t)$  和  $T_{\Sigma}^v(t) = T_{\Sigma}^v(1)/t$ , 因此  $T_{q_v}^v(t)$  可写成:

$$T_{q_v}^v(t) = T_{q_v}^v(1) \frac{2^{q_v-1} - 1}{(1+t)^{q_v-1} - 1} \quad (20)$$

$$\sim T_{q_v}^v(1) \frac{2^{q_v-1} - 1}{t^{q_v-1}} \quad (t \rightarrow \infty) \quad (20')$$

把方程 (20') 代入方程 (19), 可有:

$$\sum_{t=t_0}^{\infty} g_{q_v}(\Delta x_{t_0}) \propto \sum_{t=t_0}^{\infty} \frac{1}{t^{(q_v-1)d}} \quad (21)$$

对任意  $D$  和  $q_v = 1, 2$ , 有  $(q_v - 1)d = 1$ <sup>[15]</sup>。我们假设这对  $\forall q_v$  也适用, 因此:

$$d = \frac{1}{q_v - 1} \quad (\forall q_v, \forall D) \quad (22)$$

代入方程 (21) 可知方程 (21) 是发散的。因此, 在  $t \rightarrow \infty$  极限下, GSA 必定收敛于全局极小。如果我们把方程 (18) 和方程 (22) 代入到方程 (12), 可得到

$$g_{q_v}(\Delta x_t) = c \frac{[T_{q_v}^v(t)]^{\frac{-D}{2a - D(q_v-1)}}}{\left\{ 1 + (q_v - 1)b \frac{(\Delta x_t)^2}{[T_{q_v}^v(t)]^{\frac{2}{2a - D(q_v-1)}}} \right\}^{\frac{a}{q_v-1}}} \quad (23)$$

对  $q_v = 1$  ,方程 (23) 必须恢复方程 (15) ,因此  $b = 1$  , $a = 1$  (对任意  $D$  )。对  $q_v = 2$  ,方程 (23) 必须恢复方程 (17) ,则  $b = 1$  , $a = (D + 1)/2$  (对任意  $D$  )。为简单化 ,我们假定

$$b = 1 \quad ( \forall q_v , \forall D ) \tag{24}$$

最后 ,利用方程 (13) 确定 (23) 的归一化系数  $\alpha$  (此时 , $C$  还是剩下自由参量的函数) ,并且把方程 (24) 代入方程 (23) 有

$$g_{q_v}(\Delta x_t) = \left(\frac{q_v - 1}{\pi}\right)^{D/2} \frac{\Gamma(\frac{a}{q_v - 1})}{\Gamma(\frac{a}{q_v - 1} - \frac{D}{2})} \frac{[T_{q_v}^v]^{-\frac{D}{2a - D(q_v - 1)}}}{\left\{1 + (q_v - 1) \frac{(\Delta x_t)^2}{[T_{q_v}^v(t)]^{\frac{2}{2a - D(q_v - 1)}}}\right\}^{\frac{a}{q_v - 1}}} \tag{25}$$

此时只有一个不确定的数  $\alpha(q_v, D)$  ,但前面提到的 , $\alpha(1, D) = 1$  , $\alpha(2, D) = (D + 1)/2$  ,所以我们采用最简单的线性插值去决定与  $q_v$  的关系 :

$$a = 1 + \frac{D - 1}{2}(q_v - 1) \quad ( \forall q_v , \forall D ) \tag{26}$$

由此 ,我们得到最终的跃迁分布形式 :

$$g_{q_v}(\Delta x_t) = \left(\frac{q_v - 1}{\pi}\right)^{D/2} \frac{\Gamma(\frac{1}{q_v - 1} + \frac{D - 1}{2})}{\Gamma(\frac{1}{q_v - 1} - \frac{1}{2})} \frac{[T_{q_v}^v(t)]^{-\frac{D}{3 - q_v}}}{\left\{1 + (q_v - 1) \frac{(\Delta x_t)^2}{[T_{q_v}^v(t)]^{\frac{2}{3 - q_v}}}\right\}^{\frac{1}{q_v - 1} + \frac{D - 1}{2}}} ( \forall q_v , \forall D ) \tag{27}$$

此分布的二级矩在  $q_v \geq 5/3$  时发散 ,而当  $q_v \geq 3$  时 ,就不可归一化了。并没有特别的原因要求  $T_{q_v}^v = T_{q_v}^a$  ,但为了简单起见 ,可以让两者相等 ,以下我们用  $T$  表示。

推广模拟退火步骤如前面的模拟退火 ,只是要用推广的降温方式 ,推广的跃迁几率分布 和推广的接收几率分布。

### 4 推广模拟退火性能的测试和比较

通常 ,优化算法的特点可以通过标准测试函数反映出来。我们将通过测试四种标准函数来展示模拟退火方法的性能。这四种标准测试函数分别是 Goldstein-Price(GP)函数 ,Brain(BR)函数 ,3 维 Hartman 函数  $H_3$  和 Shubert(SH)函数。它们已经在 Dixon 和 Torn 等人的著作中作了非常仔细的讨论<sup>[20, 21]</sup>。本章我们选了七种随机优化算法同推广模拟退火作比较。这七种随机优化算法分别是纯随机搜索(PRS) ,多始点法(MS) ,模拟退火(SA1, SA2) ,禁忌搜索(TS) ,最低能轴算法(LEP)和最近邻轴算法(NNP)。

GSA 和以上七种算法的性能可以用它们的函数平均调用次数(average number of function calls over the successful runs)来衡量。对每种测试函数 ,我们执行 1000 次 GSA<sup>[16, 18]</sup> ,如果有  $K$  次成功找到全局极小 ,则平均函数调用次数就是对这  $K$  次执行的所有函数调用次数求平均。我们还统计了另一个指标 :执行成功率(percent of successful runs) ,即  $K/1000$ 。为了便于同 Serra<sup>[13]</sup>等人的轴算法进行比较 ,在测试中我们采用了他们的做法 ,即保证此时的执行成功率大于 95%。停止准则以是否找到全局极小为准 ,当

函数值和全局极小值的差异不到 3% 时,就认为已找到全局极小。如果算法执行超过了一足够大的特定步数还未找到全局极小,则认为体系已陷入了一个局部极小不可能找到全局极小,算法也停止。此特定步数被设置得足够大以确保此时体系确实陷入了一个局部极小。

表 1 四种标准测试函数中八种优化算法的平均函数调用次数

Method	GP	BR	H <sub>3</sub>	SH
PRS	5125	4850	5280	6700
MS	4400	1600	2500	—
SA1	5439	2700	3416	241215
SA2	563	505	1459	780
TS	486	492	508	727
LEP	112	144	122	281
NNP	153	68	52	114
GSA	146	62	20	104

我们的结果列在表 1 中(与 GSA 进行比较的七种方法的结果引自文献 [22, 13])。从表 1 可看出,尽管与 GSA 进行比较的其它七种方法中 NNP 是最好的,对所有四种测试函数,GSA 的平均函数调用次数都比 NNP 的小。除了在 GP 函数中 GSA 的平均函数调用次数比 LEP 略大一点外(比其它六种方法都小),在其它 3 种函数中,GSA 的结果比其它七种方法(PRS,MS,SA1,SA2,TS,LEP,NNP)都好。以上测试表明,GSA 比其它七种方法都更优越。相对于 TS 而言,GSA 的效率是 TS 的 4 到 21 倍。

5 推广模拟退火在 Thomson 问题中的应用

Thomson 问题源于物理学家 J. J. Thomson 在本世纪初提出的葡萄布丁原子模型(plum pudding model of the atoms)<sup>[23]</sup>。球面上分布的  $N$  个等量点电荷如何排布才能使体系能量最低,这就是所谓的 Thomson 问题。自该问题被提出来以后,就有 J. J. Thomson 本人和 D. Hibert 以及后来的许多数学家研究它<sup>[23]</sup>,但是解析方面的工作非常少,大部分工作都是采用数值方法。

Thomson 问题的库仑能量函数是：

$$E = \frac{1}{2} \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

(28)

Erber 等人<sup>[24, 25]</sup>指出,Thomson 问题的局域极小数目随粒子数成指数增加( $0.382 \times \exp(0.0479 N)$ ),并且包含全局极小的漏斗区域往往相对比较小,因此 Thomson 问题被认为是优化方法的一个理想的测试体系<sup>[26]</sup>。

Thomson 问题的解已经被很多优化方法如最速下降<sup>[27~29]</sup>,Monte Carlo<sup>[30, 31]</sup>,模拟退火<sup>[32]</sup>,约束全局优化法<sup>[33]</sup>,和遗传算法<sup>[26]</sup>搜索过。最速下降只能给出局域极小而不是全局极小。而 Monte Carlo 和模拟退火虽然能够通过跳到附近的极小点去改善当前的极小点,由于他们的局域跃迁方式,从一个极小跳到另一个极小要花费相当长的时间,对于大体系,这些方法也不实际。

5.1 计算结果

我们用 GSA 计算了从  $N = 2$  到  $N = 220$  的构形。对每一个  $N$  ,在很高的温度下 ,我们从一些随机的结构出发执行推广模拟退火。我们发现 ,对不大的粒子数  $N$  ,GSA 有时会陷入局域极小 ,但只需要多算几个初始态 ,全局极小就能找到。GSA 成功地找到了一些能量更低的新结构 ,表 2 列出了我们的  $N = 155 \sim 220$  的结果。我们新发现的结构在能量上比文献中结构平均低到  $10^{-2}$ 。对  $N = 156, 177, 188, 190, 197$  ,文献中的结构在我们的计算中被证明是亚稳结构。对  $N = 198$  ,除了最低能量结构外 ,我们还找到了 3 个亚稳结构 ,能量都比文献中结果要低。需要说明的是 ,为了提高结果精度 ,我们在 GSA 结束后用类似于分子动力学中的最速下降和共轭梯度法作了局域搜索。

表 2  $N = 156 \sim 120$  的能量结果。括号里面是已有的文献结果<sup>[26]</sup>

$N$	Energy	$N$	Energy	$N$	Energy
156	11092.79831146 ( 11092.80311478 )	178	14531.30955259	198	18064.26217720 ( 18064.28806296 )
		179	14698.75459423		
157	11238.90304116	180	14867.09992753	199	18251.08249564
158	11385.99018620	181	15036.46723977	200	18438.8427153
159	11534.02396096	182	15206.73061091	201	18627.59122628
160	11683.05480555	183	15378.16657103	202	18817.20471827
161	11833.08473947	184	15550.42145032	203	19007.98120458
162	11984.05033581	185	15723.72007408	204	19199.54077561
163	12136.01305322	186	15897.89743705	205	19392.36915239
164	12288.93010532	187	16072.97518632	206	19585.95585655
165	12442.80445137	188	16249.22267888	207	19780.65690934
166	12597.64907132		( 16249.25013148 )	208	19976.20326121
167	12753.46942975	189	16426.37193887	209	20172.75468067
168	12910.21267227	190	16604.42833850	210	20370.25161515
169	13068.00645113		( 16604.44596500 )	211	20568.74060279
170	13226.68107856	191	16783.45221937	212	20768.05308597
171	13386.35593072	192	16963.33838646	213	20968.61202550
172	13547.01810879	193	17144.56474088	214	21169.91041038
173	13708.63524304	194	17326.61613647	215	21372.34878935
174	13871.18709229	195	17509.48930394	216	21575.59637787
175	14034.78130694	196	17693.46054809	217	21779.85608042
176	14199.35477563	196	( 17693.46055212 )	218	21985.26394893
177	14364.83754530 ( 14364.85051922 )	197	17878.34016257	219	22191.48547482
			( 17878.38274577 )	220	22398.65560254



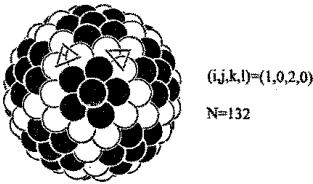


图 2  $N = 132$  的基态结构

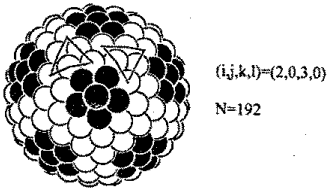


图 3  $N = 192$  的基态结构

5.2 Thomson 问题中的二十面体结构研究

对  $N = 12, 32, 72, 122$  和  $132$  ,它们的结构是二十面体结构 ,有 12 个五配位的粒子且它们分布在正二十面体的 12 个顶点上 ,剩下的  $N - 12$  个粒子是六配位的。每个五配位的粒子同它周围的五个配位粒子构成五边形(图 2 中的黑色粒子)。Altschuler 等人猜想对  $N = 10(m^2 + n^2 + mn)$  ( $m, n$  为正整数) ,全局极小就是这种高对称的二十面体结构<sup>[34]</sup>。他们用随机优化算法证明了对  $N \leq 132$  猜想都是对的。Altschuler 的公式预言在  $132 < N \leq 372$  的范围内 ,192, 212 和 272 也应该是这种二十面体结构。对  $N = 192$  ,我们得到了与 Morris 同样的全局极小。 $N = 192$  的结构确实是二十面体 ,与 Altschuler 的猜想一致(图 3)。因为以前的优化结果都限制在  $N \leq 200$  ,所以对  $N > 200$  ,Altschuler 的猜想不知道是否正确。我们用 GSA 搜索  $N > 200$  的最低能量结构( $N \leq 272$ )。对  $N = 212$  ,我们确实发现它是一个二十面体结构(图 4)。对  $N = 272$  ,有  $3 \times 10^5$  个亚稳结构 ,用一般的优化算法很难找到它的全局极小。GSA 优化得到最低能量结构如图 5 所示 ,可以看出 ,它也是一个二十面体结构。所以 ,通过用 GSA 优化 Thomson 问题的结构 ,我们把 Altschuler 猜想的证明从  $N \leq 132$  扩展到  $N \leq 272$ 。

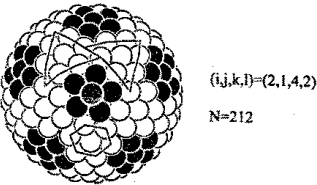


图 4  $N = 212$  的基态结构

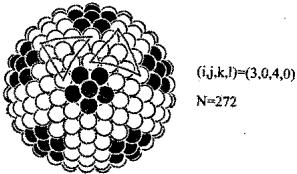


图 5  $N = 272$  的基态结构

对于以上二十面体结构 ,我们给出了一个新的拓朴模型。这个拓朴模型能够清楚地直接描述粒子在球面上的分布。我们的模型是 除了 12 个五边形外 ,剩下的  $N - 72$  的粒子在球面上构成 20 个“球面正三角形”。三角形的每一边有  $k = i + j + 1$  个粒子( $i, j$  是整数 , $i > j, i > 0, j \geq 0$ )。每一个正三角形同它最近邻正三角形 有三个最近邻三角形 有一个“球面菱形”的重叠区域(图 6)这个菱形的两边分别有  $i$  个粒子和  $j$  个粒子 ,所以菱形区域里有  $l = i \times j$  个粒子。由此可得出粒子数公式 :
$$N = 72 + 20 \left[ \frac{(i + j + 1)(i + j + 2)}{2} - \frac{3}{2}ij \right] \quad (i, j \text{ 整数 } i > j, i > 0, j \geq 0)$$
 如果我们让  $i = n + m - 3, j = n - 3$  或  $m - 3$  ,则我们的公式就转换为

Altschuler 的公式;如果让  $m = \max \{i-j, j+3\}$ ,  $n = \min \{i-j, j+3\}$  则 Altschuler 的公式就转换为我们的公式。在  $N = 132, 192, 212, 272$  的结构图中我们列出了对应的  $(i, j, k, l)$ 。事实上在  $N \leq 1000$  范围内我们证明了这个拓扑模型的正确性。

6 推广模拟退火在团簇结构研究方面的应用

近年来物理学家和化学家对团簇发生了广泛且浓厚的兴趣<sup>[1~9]</sup>。由于团簇的结构对理解和解释团簇许多奇异特性起着非常重要的作用,因此团簇结构的确定自然就成了团簇科学中最为基本的问题之一。遗憾的是,目前还没有一种能直接测量在自由状态下小团簇结构的方法。理论研究的任务是寻找团簇的稳定结构,并为实验提供真实的候选结构。寻找团簇的最稳定结构时,往往需要采用不同的优化方法,我们将应用推广模拟退火算法<sup>[19,16,35,18,36,13]</sup>研究 Ni 团簇的基态原子结构(粒子数  $n$  从 2 到 55)。Ni 团簇中原子相互作用势采用 Sutton 和 Chen 改进的长程 FS 势<sup>[37~38]</sup>:

$$V = \epsilon \sum_i \left[ \frac{1}{2} \sum_{j \neq i} \left( \frac{a}{r_{ij}} \right)^k - c \rho_i^{\frac{1}{2}} \right]$$

(29)

其中

$$\rho_i = \sum_{j \neq i} \left( \frac{a}{r_{ij}} \right)^m$$

(30)

这里  $r_{ij}$  是原子  $i$  和  $j$  之间的距离,  $a$  是晶格常数,  $c$  是无量纲参数,  $\epsilon$  是有能量量纲的参数,  $m$  和  $k$  是整数。对 Ni, 上面的参数是:  $a = 3.52 \text{ \AA}$ ,  $c = 39.432$ ,  $\epsilon = 1.5707 \times 10^{-2} \text{ eV}$ ,  $m = 6$ ,  $k = 9$ 。Nayak 等人的工作表明 FS 势产生的小的 Ni 团簇的结构同第一性原理结果一致<sup>[39]</sup>。

我们运用推广模拟退火优化 Ni 团簇的能量函数,从而找出最低能量构形。 $q_v$  和  $q_a$  被设成 2.62 和 -5。在本工作中,我们对接收几率引入新温度  $T/t$ ,  $T$  为跃迁温度。我们测试表明,这种技巧加速了算法收敛。选取初温使得在这个温度下接收率为 86%,在推广模拟退火温度降到 0.01 时停止,然后用共轭梯度法作局域搜索提高最后结果的精度。对每个  $n$ ,我们运行 1000 个随机初始态,所得到的最低能量结构作为基态结构。

图 7 给出了  $Ni_n$  ( $n = 2 \sim 55$ ) 的一些典型结构。从  $n = 9$  到 25, Ni 团簇的结构或多或少都是二十面体类型的。 $Ni_{24}$  和  $Ni_{25}$  分别可看成三重二十面体加一个原子 (triple icosahedron plus one atom) 和三重二十面体加上二个原子 (tripe icosahedron plus two atoms)(见图 7),不过这个三重二十面体扭曲得较厉害。这两个构形正好同 Riley 等人的

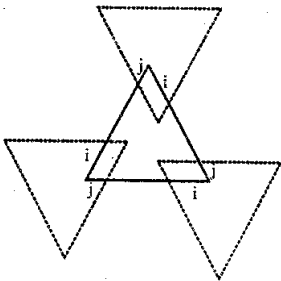


图 6 拓扑模型的示意

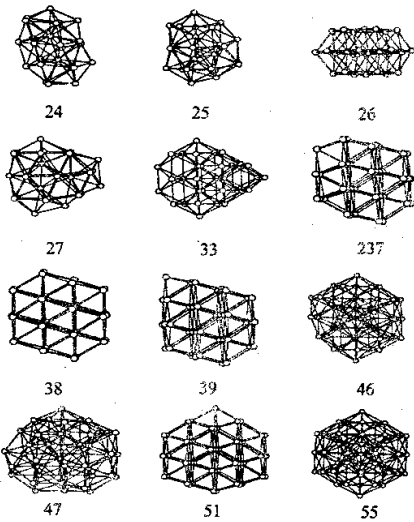


图 7 Ni 团簇的典型结构

实验一致<sup>[40]</sup>。不过,与 Riley 认为  $Ni_{26}$  是一个 Polyicosahedron 不同,我们发现  $Ni_{26}$  是一个堆积错的 fcc 的晶体碎片,它有  $D_{3h}$  的对称性。从图 7 可看到,  $Ni_{26}$  有三个 fcc(111) 面,这三个面分别有 7,12,7 个原子,如果第一层是 A,第二层是 B,则第三层还是 A,即三个面排成 ABA 的模式,所以我们可以称它为 hcp 碎片。

从  $n = 27$  到  $n = 36$  除了  $n = 33$  外,所有的结构都是无序的,但大部分结构都可以由  $n \leq 13$  的小团簇组合出来。组合的时候有一些面和边是共用的。例如,图 7 中的  $Ni_{27}$  可以由  $Ni_{10}$ ,  $Ni_{11}$  和一个六角双锥构成。

$Ni_{33}$  被认为是一个十面体。它的主要部分是一个五角双锥,沿五角双锥的五重轴有 4 个原子(两个锥顶各一个,轴中部有两个)。另有 14 个原子附在这个主要部分周围。图 7 可看出,  $Ni_{33}$  结构的右边有两个完整的 fcc(111) 面形成(即指完整的六角形)。在  $n = 37$  到 39 范围内,我们发现 fcc 结构对能量更有利。这些 fcc 结构的存在说明以前通常认为在不很大的体系中团簇结构总是采取 icosahedral 类型结构的观点是不对的。 $Ni_{37}$  的结构是一个有堆积错的 fcc 晶体碎片,它的四个 fcc(111) 面的排列是 ABCB,从左边看,ABC 是典型的 fcc 排序,从右边看,BCB,又是 hcp 结构,所以也可以认为是 fcc, hcp 的混合物。 $Ni_{38}$  被发现是一个完整的截角八面体(truncated octahedron),这与 Riley 等人最近的关于  $Ni_{38}$  的实验非常一致。同  $Ni_{38}$  相似,  $Ni_{39}$  也是一个 fcc 碎片。事实上,  $Ni_{38}$  加上一个原子就成了  $Ni_{39}$ (如图 7)。

从  $n = 40$  到  $n = 51$ ,二十面体类型,十面体类型和 fcc 碎片交替出现。例如  $Ni_{46}$ ,  $Ni_{47}$  是二十面体类型的,  $Ni_{45}$ ,  $Ni_{48}$  是十面体类型的,与  $Ni_{37}$  相似,  $Ni_{50}$  和  $Ni_{51}$  的结构是有堆积错的 fcc 碎片。

从  $n = 52$  到  $n = 55$ ,团簇的结构都是二十面体结构,直到  $N = 55$  一个完整的二十面体构成。

从上面结构的讨论可以看出,  $n = 13$  附近( $n = 9 \sim 25$ )和  $n = 55$  附近( $n = 52 \sim 55$ ),团簇的结构是二十面体类型的。在  $n = 38$  附近, fcc 类型出现。这一点是同 Riley 的实验一致的<sup>[41]</sup>。然而在这三个区域之间,团簇表现出复杂的结构。我们的结果说明通常认为在不太大尺寸下团簇总具有二十面体类型结构的观点对 Ni 团簇是不对的。

## 7 推广模拟退火的效率

### 7.1 GSA 与 CSA 和 FSA 的比较

本小节我们把推广模拟退火(GSA)同经典模拟退火(CSA)和快速模拟退火(FSA)作比较<sup>[18]</sup>。为了加快 GSA 收敛,我们让  $q_a$  与步数成线性下降。我们选取  $N = 12$  的 Thomson 问题作为测试体

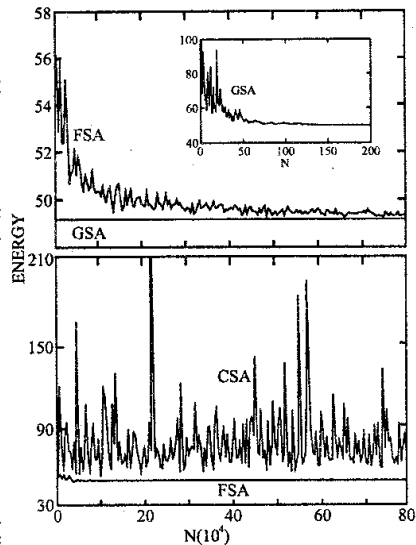


图 8 对  $N = 12$ , GSA, FSA 和 CSA 的能量随时间的收敛

系,将 GSA ,CSA ,FSA 分别用到这个体系中。GSA ,CSA ,FSA 都从相同的初始结构出发 ,且运行同样的步数。图 8 给出了  $N = 12$  的结果。为了更清楚起见 ,在图 8 中的上部分和下部分用了不同的标度。从图 8 可看出 ,GSA 仅仅需要几百步就收敛到全局极小 ,并且振荡非常小。与其他研究者展示的一样 ,图 8 中 CSA 的振荡很大 ,且收敛很慢。相对于 CSA 而言 ,FSA 的振荡急剧减小 ,收敛的速度比 CSA 更快 ,但仍比 GSA 慢得多 ,事实上直到近八十万步后它才接近全局极小 ,但还未完全收敛下来。在这个例子中 GSA 比 FSA 快了 100 倍以上。这些比较说明 GSA 以比 FSA 和 CSA 更高的几率收敛于全局极小。使用某种程度的扭曲的 Cauchy-Lorentz 跃迁分布的优势是不但局域搜索的能力保留下来 ,而且拥有比 FSA 更频繁的长距离跳跃的能力使 GSA 比 FSA 和 CSA 更快的跳出局域极小。更快的跳出局域极小意味着有更大的几率找到全局极小。

7.2 推广模拟退火效率随体系复杂性的变化

一个算法好不好 ,主要看它对复杂体系的效率如何。显然随着目标函数的变量增多 ,目标函数变得越复杂。有的算法在目标函数变量数目比较小时 ,比其他算法优越 ,但随着变量数目增大 ,它的优势越来越小 ,超过某一变量数目时还会比其他算法差。例如 :Serra 等人把轴算法 (Pivot method)<sup>[13,42,43]</sup>应用到 Lennard-Jones 团簇时 ,发现当  $N < 6$  时 ,遗传算法更有效一些 ,而当粒子数变大时 ,轴算法更有效 ,并且随着粒子数的增加这种优势越来越大。

从我们在 Thomson 问题中的研究可以看到 ,对推广模拟退火而言 ,不仅能量的振荡显著减小 ,而且收敛也比经典模拟退火和快速模拟退火快得多<sup>[18]</sup>。随着目标函数变量数目的增加 ,GSA 相对于 CSA 和 FSA 的优势是如何变化呢 ? 下面两节我们将分别在 Thomson 模型和  $N_i$  团簇中研究这个问题。

我们用  $GSA^{[16\sim 17]}$  , $CSA^{[11]}$ 和  $FSA^{[15]}$ 搜索目标函数(能量函数)的全局极小。在 GSA 中 ,对  $q_a$  我们用 Tsallis 建议的值 - 5 ,对  $T_{q_v}$  我们仍然用以前在 Thomson 问题中使用的值 2.62。为了加速收敛 ,我们引入接收温度  $T_{q_a}$  ( acceptance temperature ) ,它等于跃迁温度  $T_{q_v}$  ( visiting temperature )除以计算机时间  $t$ 。对 GSA ,FSA ,CSA 三种算法 ,我们都从同样的初始温度 ,同样的初始态(初始结构)出发 ,且在同样的跃迁温度(0.01)下停止算法。然后用共轭梯度法作更精确的局域搜索 ,这样我们就能准确地知道是否收敛到了全局极小 ,或是陷入一个局域极小。在共轭梯度法作局域搜索时 ,如果连续前后 15 步的能量差小于  $10^{-8}$  ,我们就认为已经足够精确地收敛到极小 ,这时就停止局部搜索。为了保证结果有良好的统计性 ,我们选择了足够多的初始态以保证对每种方法全局极小被找到最少 200 次。

定义  $P$  为绝对效率 ,即从高温降到低温时每试探跳跃一次找到全局极小的可能性。则  $P_{GSA}$  , $P_{FSA}$  , $P_{CSA}$  分别是用 GSA ,FSA ,CSA 通过一次跳跃找到全局极小的数目。GSA 相对于 FSA 的相对效率定义为  $P_{gf}^* = \frac{P_{GSA}}{P_{FSA}}$  ,GSA 相对于 CSA 的相对效率定义为  $P_{gc}^* = \frac{P_{GSA}}{P_{CSA}}$ 。D.J. Wales 等人认为 ,一般而言复杂体系的局域极小的数目随着能量函数的变量

数目作指数增加。局域极小的数目越大,绝对几率  $P$  的值就越小。绝对几率  $P$  可写成

$$P = \alpha e^{-\beta N} \tag{31}$$

其中  $N$  是目标函数的变量数目。虽然  $\alpha$  和  $\beta$  都对绝对几率  $P$  有贡献,但当  $N$  足够大时,  $\beta$  就起决定性的作用。因此,我们最感兴趣的是  $\beta$ 。 $\beta$  被定义为成效率因子( efficiency factor ),一个小的效率因子意味着更有效( 注意这是因为  $P$  关于  $\beta$  是减函数 )。根据以上方程 (28),GSA 相对于 FSA 和 CSA 的相对效率可以表示为:

$$P_{gf}^* = \lambda_{gf} e^{\gamma_{gf} N} \tag{32}$$

$$P_{gc}^* = \lambda_{gc} e^{\gamma_{gc} N} \tag{33}$$

其中相对效率因子  $\gamma_{gf} = \beta_{FSA} - \beta_{GSA}$  和  $\gamma_{gc} = \beta_{CSA} - \beta_{GSA}$  是两种方法的效率因子差,  $\lambda_{gf} = \frac{\alpha_{GSA}}{\alpha_{FSA}}$ ,  $\lambda_{gc} = \frac{\alpha_{GSA}}{\alpha_{CSA}}$ 。如果相对效率  $P_{gf}^*$  或者  $P_{gc}^*$  比 1 大,就意味着 GSA 比 FSA 或者 CSA 更有效。实际上我们以前的结果已经证实了在所研究的尺寸范围内 GSA 与 FSA 和 CSA 更有效。对相对效率起决定性的量是相对效率因子  $\gamma_{gf}$  ( $\gamma_{gc}$ )。如果  $\gamma_{gf}$  ( $\gamma_{gc}$ ) 大于零,那么随着目标函数变量数目的增大,相对于 FSA 和 CSA, GSA 将越来越有效。所以决定 GSA 效率的关键是看相对效率因子  $\gamma_{gf}$  ( $\gamma_{gc}$ ) 是否大于零。

7.2.1 在 Thomson 问题中的测试

如前面介绍, Thomson 问题也就是求  $N$  个等量点电荷在单位球面上如何排布使体系库仑能量最低的问题。我们在 Thomson 模型中的结果是相对效率因子  $\gamma_{gf}$  ( $\gamma_{gc}$ ) 比零大,

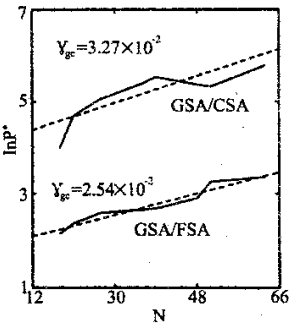


图 10 在 Ni 团簇中 GSA 相对于 FSA 和 CSA 的效率随变量数(三倍于粒子数)的变化图。拟合的直线(虚线)说明 GSA 相对于 FSA 和 CSA 的相对效率随变量数  $N$  近似于指数增加

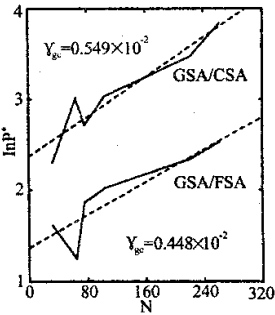


图 9 Thomson 模型中 GSA 相对于 FSA 和 CSA 的效率随变量数(两倍于粒子数)的变化图。拟合的直线(虚线)说明 GSA 相对于 FSA 和 CSA 的相对效率随变量数  $N$  近似于指数增加

即相对于 CSA 和 FSA,变量数目越大,GSA 越有效。图 9 给出了相对效率随目标函数变量数目  $N$  的变化关系。可以看出,  $\ln P^*$  随  $N$  近似线性增加。通过拟合,我们有  $\gamma_{gf} = 0.448 \times 10^{-2}$ ,  $\gamma_{gc} = 0.549 \times 10^{-2}$ 。从图 9 还可以看出  $\ln P^*$  总是比 1 大。所以 GSA 相对 FSA 和 CSA 的优势不仅表现在  $\beta$  上,而且表现在前面的系数  $\alpha$  上。正如前面提到的,目标函数的变量越多,体系就越复杂,所以随变量数目增加而增加的相对效率就意味着:相对于 CSA 和 FSA,体系越复杂,GSA 就越有效。

7.2.2 在 Ni 团簇体系中的测试

本节我们将在 Ni 团簇体系中研究 GSA 效率的变化。Ni 原子间相互作用仍用 FS 势的 Sutton-Chen 版本<sup>[38]</sup>描述(见前面 GSA 在 Ni 团簇中的应用):

我们知道 Ni 团簇体系比 Thomson 模型更复杂,找到一

个全局极小要使用更多的初始态。GSA 相对于 CSA 和 FSA 的相对效率随变量数目  $N$  的变化如图 10 所示。可以看出,随着变量数目  $N$  的增大,相对效率  $P^*$  增加。我们用直线拟合  $\ln P^* \sim N$  得到的  $\gamma_{gf}$  和  $\gamma_{gc}$  是:  $\gamma_{gf} = 2.54 \times 10^{-2}$ ,  $\gamma_{gc} = 3.27 \times 10^{-2}$ 。因此我们用 GSA 搜索 Ni 团簇结构的结论是: GSA 相对于 FSA 和 CSA 的相对效率随目标函数变量数目呈指数增加。同 Thomson 问题中一样,  $\alpha_{GSA}$  也比  $\alpha_{FSA}$  ( $\alpha_{CSA}$ ) 大。

Ni 团簇的相对效率因子  $\gamma$  比 Thomson 问题中的  $\gamma$  大,这也就是说体系越复杂, GSA 相对于 FSA 和 CSA 的优势越大。因为包含着多体项, Ni 团簇的多体能量函数比 Thomson 问题的两体库仑势要更复杂, 所以优化寻找 Ni 簇的最低能量结构比搜索 Thomson 问题的最低能量结构要更困难。比较图 9 和图 10, 我们发现无论是 Ni 团簇的相对效率  $P^*$  还是 Ni 团簇的相对效率因子  $\gamma$  都比 Thomson 问题中的  $P^*$  和  $\gamma$  大, 这说明了在复杂体系中 GSA 相对于 FSA 和 CSA 的优势更大。

### 7.2.3 讨论

为什么体系越复杂, GSA 相对于 FSA 和 CSA 的优势越大? 为了回答这个问题, 我们给出这三种算法在低温下对二维相空间的搜索图(图 11)<sup>[13]</sup>。在这个图中, 三种算法的搜索特性看得很清楚。CSA 用的是 Gaussian 分布, 这是一个典型的局域分布, 因为大部分搜索都被限制在相空间的一个局域部分(图 11 上面子图)。FSA 用 Cauchy-Lorentz 分布(图 11 的中间子图), 这是一个半局域分布, 其半局域特性使体系偶尔能从一个区域作长距离跳跃到另一个远的区域, 因此 FSA 搜索相空间比 CSA 更有效, 但仍然陷入一些区域很多时间(图 11 的中间子图的较黑的区域)。GSA 使用一个扭曲的 Cauchy-Lorentz 分布, 即 Tsallis-stariolo 跃迁分布形式, 这是一个非局域分布。它的非局域性使体系即使在低温下也有大的几率作长距离跳跃, 因此 GSA 能充分的搜索相空间(图 11 的下子图), 从而 GSA 能以更大的几率找到全局极小。在复杂体系中非局域搜索是很重要的。随着目标函数的变量数目增加, 体系复杂性也增加, 局域极小的数目增加得非常快, 则局域搜索算法更有可能陷入局域极小。GSA 是非局域搜索, 即使在低温下也有较大机会作长距离跳跃, 那么它从局域极小中跳出的可能性也大得多。这可能为什么 GSA 相对于 CSA 和 FSA 的相对效率  $P^*$  随目标函数变量数目(体系复杂性)增加的原因之一。

## 8 结 语

推广模拟退火方法是一种很高效的优化算法, 除了在物理中的应用外, 在量子化学和分子生物学中也得到了越来越多的应用。我们认为它也将在自旋玻璃等更复杂的体系中显示出它的威力。在方法发展方面, 我们相信推广模拟退火与决定性的优化算法的有机结合将进一步提高它的效率, 当然, 这需要比较细致的研究。

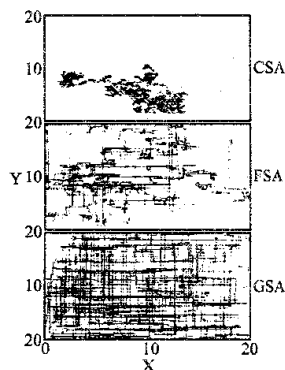


图 11 低温下对二维相空间的搜索。CSA 用 Gaussian 分布(上小图), FSA 用 Cauchy-Lorentz 分布(中间小图), GSA 用 Cauchy-Lorentz 分布的 Tsallis-Stariolo 形式。GSA 充分搜索相空间。

## 参考文献

- [ 1 ] 龚新高. 中国科学院固体所博士论文. 1993.
- [ 2 ] 冯端, 金国钧. 凝聚态物理新论. 上海科学技术出版社, 1992.
- [ 3 ] Brack M. Rev. Mod. Phys. 1993, 65 :677.
- [ 4 ] De Heer W A. Rev. Mod. Phys. 1993, 65 :611.
- [ 5 ] Tycko R, Dabbagh G, Fleming R M, Haddon R C, Makhija A V and Zahurak S M. Science. 1991, 253 : 884.
- [ 6 ] Holczer K, Klein O, Hung S M, et.al. Science. 1991, 252 :1154.
- [ 7 ] Shi Gider J S, Babcock K, Awschalom, Science. 1996, 271 :937.
- [ 8 ] Vossmeier T, Katsikas L, Giersig M, Popovic I G, Weller H. J. Phys. Chem, 1994, 98 :7665.
- [ 9 ] Goldstein A N, Echer C M, Alivisatos A P. Science, 1982, 256 :1425.
- [ 10 ] 张开明, 顾昌鑫编著. 计算物理学. 复旦大学出版社, 1987.
- [ 11 ] Kirkpatrick S, Gelatt G D Jr, and Vecchi M P. Science. 1983, 220 :671.
- [ 12 ] 康立山, 谢云, 尤矢勇, 罗祖华. 非数值并行算法 – 模拟退火算法. 科学出版社, 1994.
- [ 13 ] Serra P, Stanton A F, Kais S and Bleil R E. J. Chem. Phys. , 1997, 106 :7170.
- [ 14 ] Geman S and Geman D. IEEE Trans. Patt. Anan. Mach. Int. 1984, PAMI – 6 ( N<sup>0</sup>6 ), 721.
- [ 15 ] Szu H and Hartley R. Phys. Lett. , 1987, A122 :157.
- [ 16 ] Tsallis C, Stariolo D A. Physica, 1996, A223 :395.
- [ 17 ] Ioan Addricioaei and John E. Straub. Phys. Rev. , 1996, E53 :3055.
- [ 18 ] Xiang Y, Sun D Y, Fan W, Gong X G. Phys. Lett, 1997, A233 :216.
- [ 19 ] Tsallis C J. Stat. Phys, 1988, 52 :479.
- [ 20 ] Dixon L C W, Szego G P. Towards Global Optimization ( North-Holland, New York, 1975 ), vol. 1.
- [ 21 ] Torn A, Zilniskas A. Global Optimization ( Springer-Verlag, Berlin, 1989 ).
- [ 22 ] Cvijovic D, Klinowski J, Science. 1995, 267 :664.
- [ 23 ] Whyte L L. Am. math. Mthly, 1952, 59 :606.
- [ 24 ] Erber T and Hockney C M. Phys. Rev. Lett, 1995, 74 :1482.
- [ 25 ] Erber T and Hockney G M. ( unpublished ).
- [ 26 ] Morris J R, Deaven D M and Ho K M. Phys. Rev. , 1996, B53 :R1740.
- [ 27 ] Anderssen R S. in Optimisation, Univ of Queensland Press, st. Lucia, Australia ( 1972 ).
- [ 28 ] Dekkers A, Aarts E. Math. Program. 1991, 50 :3.
- [ 29 ] Erber T and Hockney G M. J. Phys. , 1991, A24 :L1369.
- [ 30 ] Glasser L and Every A G. J. Phys. , 1992, A25 :2473.
- [ 31 ] Webb S. Nature, 1986, 323 :20.
- [ 32 ] Wille L T. Nature, 1986, 324 :46.
- [ 33 ] Altschuler E L, Williams T J, Ratner E R, Dowla F and Wooten. F. Phys. Rev. Lett. , 1994, 72 : 2671 ; 1995, 74 :1483.
- [ 34 ] Altschuler E L, Williams T J, Rater E R, Tipton R, Stong R, Dowla F, Wooten F. Phys. Rev. Lett, 1997, 78 :2681.
- [ 35 ] Hansmann E U H. Phys. 1997, A242 :250.
- [ 36 ] Xiang Y, Sun D Y, Gong X G. J. Phys. Chem, 2000, A104 :2746.

- [ 37 ] Finnis M W , Sinclair J E. Philos. Magl. 1984 ,50 :45.
- [ 38 ] Sutton A P , Chen J. Philos. Mag. Lett. ,1990 ,61 :139.
- [ 39 ] Doye J P K , Wales D J , New J. Chem. ,1998 ,22 :733 ;Nayak S K , Khanna S N , Rao B K , Jena P J. Phys. Chem. , 1997 , A101 :1072. Reddy B V , Nayak S K , Khanna S N , Rao B K , Jena P J. Phys. Chem. ,1998 , A102 :1748.
- [ 40 ] Parks E K , Zhu L , Ho J , Riley S J. J. Chem. Phys. ,1995 ,102 :7377.
- [ 41 ] Parks E K , Niemann , G C Kerns K P , Riley S J. J. Chem. Phys ,1997 ,107 :1861.
- [ 42 ] Stanton A F , Bieil R E , Kais S J. Comp. Chem ,1997 ,18 :594.
- [ 43 ] Serra P , Stanton A F , Kais S. Phys. Rev. , 1997 E55 ;Phys. Rev. ,1997 ,E55 :1162.

## GENERALIZED SIMULATED ANNEALING METHOD AND ITS APPLICATION

Xiang Yang and Gong Xingao

( *Institute of Solid State Physics ,Chinese Academy of Sciences 230031 Hefei ,China.* )

**Abstract :** After briefly reviewed the basic idea of the generalized simulated annealing ( GSA ) method ,we have explored its efficiency comparing to other well known methods. We find that GSA is more efficient than other simulating annealing methods. some results on the atomic structure of the cluster and the Thomson model will be also presented.

**Keywords :** optimization , generalized simulated annealing method , Thomas model