

doi: 10.3969/j.issn.1000-8349.2016.03.04

MCMC 方法在星系形成半解析模型中的应用

李仕杰^{1,2}, 杨小虎³

(1. 中国科学院 上海天文台 星系与宇宙学重点实验室, 上海 200030; 2. 中国科学院大学, 北京 100049;
3. 上海交通大学 天文与天体物理研究中心, 上海 200240)

摘要: 星系形成的半解析模型是理解星系形成中的重子物理过程的重要方法, 但存在显著的缺点: 物理参数太多, 调控过程复杂。MCMC (Markov chain Monte Carlo) 方法是现代统计计算中最重要的算法之一, 通过 MCMC 方法可以得到星系形成半解析模型中众多物理参数的有效范围。简要介绍了半解析模型的主要物理过程和 MCMC 方法, 综述了近年来 MCMC 方法在星系形成半解析模型中的应用和成果。这些结果表明 MCMC 方法对于限制半解析模型有很好的作用, 对更好地理解星系形成中的物理过程起到促进作用。

关键词: 星系形成; 半解析模型; MCMC 方法

中图分类号: P157.9 **文献标识码:** A

1 引 言

在标准冷暗物质晕宇宙学模型中, 宇宙结构形成是一个由下到上 (bottom-up) 的等级成团过程。现代标准星系形成理论认为, 星系是暗晕内热气体冷却坍缩的产物^[1]。因此人们相信, 重子物理过程跟暗晕的性质有着密切的联系。星系形成的半解析模型就是通过暗晕并合树结合各种重子物理过程来研究星系的形成与演化^[2]。

半解析模型以数值模拟或者蒙特卡罗方法得到的暗晕并合树为基础, 结合诸如气体冷却、恒星形成、超新星反馈、活动星系核反馈、星系并合等众多重子物理过程, 得到星系的演化历史以及众多性质特征 (如质量、光度、冷气体质量、元素丰度、黑洞质量等)。它以暗晕并合树为输入, 在并合树的最高端也就是最开始红移处, 假设每个暗晕内都有一定质量的热气体, 热气体的质量由暗晕质量以及宇宙学模型确定。之后随着并合树向下也就是红移由高到

收稿日期: 2016-03-28; 修回日期: 2016-05-25

资助项目: 973 项目 (2015 CB857002)

通讯作者: 李仕杰, sjli@shao.ac.cn

低, 热气体会慢慢冷却坍缩形成恒星, 恒星聚在一起就形成了星系。随着时间演化, 恒星会慢慢死亡, 超新星爆发会将冷气体加热进而导致恒星形成减缓, 而活动星系核所发出的辐射也能对暗晕中气体下落和冷却起到抑制作用。根据星系动力学, 由于粘滞耗散的作用, 卫星星系会并合进中心星系之中。还有其他物理过程, 这里不一一赘述。通过这些重子物理过程最终可以得到星系的诸多性质, 进而得到星系的一些统计特征如质量函数、光度函数等。通过比较模型得到的星系统计特征与观测的异同, 然后再调节相应的重子物理过程中的参数, 直到得到与观测比较相符的结果。

White 和 Frenk^[3] 于 1991 年首次提出了半解析模型的基本框架。之后经过 Kauffmann 等诸学者的努力^[4-13], 半解析模型得到了长足的发展。但是, 半解析模型也有一个显著的缺点——参数太多。因为半解析模型用到了星系形成过程中的众多重子物理过程, 而我们现在对这些过程的了解还不够深入, 只能用唯象的模型去描述这些过程, 这就导致了半解析模型中可调的参数太多。在很长一段时间里, 半解析模型的研究者只能手动调节这些参数, 然后用肉眼比对模型结果与观测的异同。因涉及到众多物理参数, 导致这个过程很复杂并且耗费时间。直到 2008 年, Kampakoglou 等人^[14] 将 MCMC (Markov Chain Monte Carlo) 方法应用到了半解析模型中, 大大提高了半解析模型的效率。

蒙特卡罗方法^[15] 是一种根据给定概率分布, 产生随机数, 利用随机数进行数值模拟的方法。而 MCMC 方法^[16] 即马尔科夫链蒙特卡罗方法则是一种基于马尔科夫链的采样机制, 它能自动搜索概率分布比较大的区域, 所以能把更多的计算资源用在关键区域。它的基本思想是通过一个平稳分布为 $\pi(x)$ 的马尔科夫链来得到 $\pi(x)$ 的样本, 基于这些样本就可以做各种统计推断。MCMC 方法应用到半解析模型中, 可以更快更准确地搜索重子物理过程的参数空间, 将以前手动调节变为自动搜索, 大大提高了半解析模型的效率。

本文结构安排如下: 第 2 章介绍了半解析模型中的物理过程; 第 3 章介绍了 MCMC 方法如何应用到半解析模型, 并列举了近年来 MCMC 方法结合半解析模型得到的一些成果; 最后在第 4 章给出了总结。

2 半解析模型中的物理过程

半解析模型中涉及到重子相关的物理过程很多, 因为对这些过程的了解较少, 所以现在一般都是用简单的经验公式来描述。半解析模型研究最著名的研究小组有两个: Munich group^[13] 和 Durham group^[11], 当然还有其他研究组, 如 SantaCruz group^[17], Kang 等人^[9], Monaco 等人^[12], Nagashima 等人^[18]。因篇幅所限, 本节所叙述的物理过程经验公式以 Munich group 为主。

2.1 气体冷却

现代标准的暗物质晕中的星系形成理论认为, 在暗晕坍缩过程中所产生的激波会将其中的气体加热至位力温度。根据 White 和 Frenk^[3] 的理论, 暗晕中的热气体是等温球对称径向分布的, 径向轮廓为:

$$\rho_g(r) = \frac{m_{\text{hot}}}{4\pi r_{\text{vir}} r^2}, \quad (1)$$

其中 m_{hot} 是暗晕内所包含的热气体总质量。Kauffmann 等人^[19] 引入冷却半径, 定义为此处的冷却时间等于暗晕的动力学时标 $t_{\text{dyn}} = R_{\text{vir}}/V_{\text{vir}}$; 而 Kang 等人^[9] 则将此时间取为宇宙的年龄。两种定义各有道理, 现在的研究并不能确定这两种定义孰优孰劣。当冷却半径小于位力半径时, 热气体以准静态方式冷却, 在给定时长 Δt 里, 冷却的气体质量为:

$$\Delta m_{\text{cool}} = 0.5 m_{\text{hot}} r_{\text{cool}} v_{\text{vir}} R_{\text{vir}}^{-2} \Delta t. \quad (2)$$

当冷却半径大于位力半径时, 热气体达不到热平衡状态, 冷却率主要由热气体吸积到中心星系的吸积率决定, 可以近似地表示为:

$$\frac{m_{\text{cool}}}{dt} = \frac{m_{\text{hot}}}{t_{\text{cool}}}. \quad (3)$$

这种冷却方式称为快速冷却。半解析模型中, 气体冷却过程一般没有自由参数。

2.2 恒星形成

恒星形成率一般写成以下形式:

$$\psi = \alpha m_{\text{cold}}/t_{\star}, \quad (4)$$

其中, α 是恒星形成效率, 为模型自由参数, m_{cold} 为冷气体质量, t_{\star} 为恒星形成时标。不同作者采用的 t_{\star} 也不相同。

Munich group^[5] 取星系的动力学时标为 t_{\star} , 即 $t_{\star} = t_{\text{dyn}} = 3r_{\text{d}}/v_{\text{vir}}$, 其中, r_{d} 为星系的标长。在 Munich 模型中^[20], 恒星形成存在一个临界面密度 Σ_{crit} , 因此形成恒星的气体有一个临界质量, 表示为:

$$m_{\text{crit}} = 9\pi r_{\text{d}}^2 \Sigma_{\text{crit}} = 5.7 \times 10^6 M_{\odot} \frac{v_{\text{vir}}}{\text{km} \cdot \text{s}^{-1}} \left(\frac{r_{\text{d}}}{\text{kpc}} \right). \quad (5)$$

因此, 式 (4) 可以写为:

$$\Delta M_{\star} = \begin{cases} 0 & (m_{\text{gas}} < m_{\text{crit}}) \\ \alpha(m_{\text{cold}} - m_{\text{crit}})/t_{\text{dyn}} \Delta t & (m_{\text{gas}} \geq m_{\text{crit}}) \end{cases}. \quad (6)$$

即当气体质量小于临界质量时, 恒星形成过程停止; 当气体质量大于临界质量时, 超出部分的气体质量转化为恒星质量。在这个过程中, α 是自由参数。

2.3 星系的并合与星暴过程

在等级成团模型中, 大暗晕是由小暗晕并合而来。当小暗晕进入到大暗晕中成为其一部分时, 位于暗晕中心的星系也会受到影响, 小暗晕中的星系变为大暗晕中星系的卫星星系。根据 Binney 等人^[21] 的动力学模型, 动力学耗散时标 t_{fric} 可以表示为:

$$t_{\text{fric}} = 1.17 \frac{v_{\text{vir}} r_{\text{sat}}^2}{G m_{\text{sat}} \ln \Lambda}, \quad (7)$$

其中 r_{sat} 代表最后追踪到卫星星系暗晕时其中心到中心星系的距离, m_{sat} 指卫星星系所包含的恒星与冷气体总质量, 哥伦布势 $\ln \Lambda = \ln(1 + m_{\text{vir}}/m_{\text{sat}})$, 其中 m_{vir} 和 m_{sat} 分别是暗晕和小暗晕的位力质量。当卫星星系经过动力学耗散时标 t_{fric} 后, 会和中心星系并合^[22]。在并合过程中, 会发生星暴过程, 产生大规模的恒星形成, 并导致星系中心黑洞形成和增长。对于星暴过程, Munich 模型采用 Somerville 等人^[23] 的碰撞星暴模型, 在一次星暴过程中, 形成的恒星质量为:

$$\Delta m_* = (m_{\text{gas,sat}} + m_{\text{gas,cen}}) \beta_{\text{burst}} (m_{\text{sat}}/m_{\text{cen}})^{\alpha_{\text{burst}}}, \quad (8)$$

其中, $(m_{\text{gas,sat}} + m_{\text{gas,cen}})$ 为两星系的冷气体总质量, $(m_{\text{sat}}/m_{\text{cen}})$ 为两星系恒星与冷气体质量和的比值, α_{burst} 和 β_{burst} 为参数, 其值可以由数值模拟得到^[24]。

星系的并合除了影响核球, 也会对中心星系的黑洞质量产生影响。并合时, 卫星星系的黑洞质量直接加入到中心星系的黑洞质量之中。此外, 还会有一部分冷气体在并合过程中落入中心星系黑洞而导致中心星系黑洞质量变大, 进而靠吸积产生的射电辐射影响之后的气体冷却。这种黑洞增长模式被称为“类星体模式”(quasar mode), 在这种模式下, 黑洞增长的质量为:

$$\Delta m_{\text{quasar}} = \frac{f_{\text{BH}}(m_{\text{sat}}/m_{\text{cen}})(m_{\text{gas,sat}} + m_{\text{gas,cen}})}{1 + (280 \text{ km} \cdot \text{s}^{-1}/v_{\text{vir}})^2}, \quad (9)$$

其中 $(m_{\text{sat}}/m_{\text{cen}})$ 和 $(m_{\text{gas,sat}} + m_{\text{gas,cen}})$ 与式 (8) 中含义相同, f_{BH} 为自由参数。

2.4 超新星反馈

随着时间演化, 恒星进入末期, 大质量恒星会成为超新星。超新星爆发会产生大量能量, 进而加热和抛射星系中的冷气体, 这个过程即为超新星反馈。对于给定的恒星初始质量函数, 单位时间内超新星爆发所释放的能量正比于新形成恒星的总质量:

$$\Delta E_{\text{SN}} = 0.5 \epsilon_{\text{halo}} \Delta m_* v_{\text{SN}}^2, \quad (10)$$

其中, $0.5 v_{\text{SN}}^2$ 为单位新形成恒星超新星爆发后所抛射的能量, v_{SN} 为超新星抛射出来物体的速度, ϵ_{halo} 为超新星爆发所产生能量中用于加热和抛射气体的比例。 v_{SN} 和 ϵ_{halo} 都是自由参数。超新星爆发所能加热的气体质量正比于超新星质量, 也就正比于新形成恒星质量:

$$\Delta m_{\text{reheat}} = \epsilon_{\text{disk}} \Delta m_*, \quad (11)$$

其中, ϵ_{disk} 是自由参数。

如果超新星爆发的能量用来加热气体之后还有多余, 这些多余的能量可以用来加热暗晕中的部分热气体, 把他们抛射出暗晕的引力势阱。这部分被抛射出去的热气体质量为:

$$\Delta m_{\text{eject}} = \left(\epsilon_{\text{halo}} \frac{v_{\text{SN}}^2}{v_{\text{vir}}^2} - \epsilon_{\text{disk}} \right) \Delta m_*. \quad (12)$$

因此, 在 Munich 模型中, 热气体分为晕中热气体 m_{hot} 和抛射出去的热气体 m_{ejected} 两类。注意, 这里用的是 m_{ejected} 而不是 m_{eject} , m_{ejected} 代表被抛射出去的热气体的总质量。这些

被抛射出去的热气体不会永远离开暗晕，随着暗晕的增长，这部分气体会慢慢回到暗晕中，这个过程被称为再结合 (reincorporation)。再结合的热气体质量为^[8]：

$$\Delta m_{\text{reincorporate}} = \gamma_{\text{ej}}(m_{\text{ejected}}/t_{\text{dyn}}^{\text{halo}})\Delta t \quad , \quad (13)$$

其中， γ_{ej} 是暗晕的单位动力学时间 $t_{\text{dyn}}^{\text{halo}} = r_{\text{vir}}/v_{\text{vir}}$ 内被再结合的气体的质量比，它是一个自由参数。

2.5 活动星系核反馈

在星系的中心，活动星系核所发出的辐射能对晕中气体的下落和冷却起到抑制作用，进而影响恒星的形成，这个过程被称为活动星系核的反馈作用。在热气体下落时，一部分进入星系中心的黑洞，使得黑洞质量增长，由于这种持续不断的黑洞吸积，使得活动星系核产生能量较低的射电波段辐射，这被称为黑洞增长的射电模式 (radio mode)。黑洞增长的另一种模式——类星体模式，已经在前面介绍过。射电模式黑洞增长率可以写为^[10]：

$$\dot{m}_{\text{radio}} = \kappa_{\text{AGN}} \left(\frac{m_{\text{BH}}}{10^8 M_{\odot}} \right) \left(\frac{f_{\text{hot}}}{0.1} \right) \left(\frac{v_{\text{vir}}}{200 \text{ km} \cdot \text{s}^{-1}} \right) \quad , \quad (14)$$

其中， m_{BH} 是黑洞质量， $f_{\text{hot}} = m_{\text{hot}}/M_{\text{vir}}$ 是热气体质量占整个暗晕质量的比例， κ_{AGN} 是一个描述吸积率的自由参数。在射电模式中，活动星系核可以提供足够多的能量来加热星际介质，从而抑制气体的冷却过程。活动星系核反馈的产能率为：

$$L_{\text{BH}} = \eta \dot{m}_{\text{radio}} c^2 \quad , \quad (15)$$

其中， c 是光速， $\eta = 0.1$ 为标准的质能转换效率。因部分热气体被中心黑洞吸积，因此气体冷却质量修正为：

$$m'_{\text{cool}} = m_{\text{cool}} - 2L_{\text{BH}}/v_{\text{vir}}^2 \Delta t \quad , \quad (16)$$

如果 $m_{\text{cool}} < 2L_{\text{BH}}/v_{\text{vir}}^2 \Delta t$ ， m'_{cool} 设为 0。由式 (15) 和式 (16) 可以看出，活动星系核的加热效率与黑洞质量成正比，即质量越大，加热效率越高。

2.6 化学演化及基本方程

在 Munich 模型中^[25-27]，根据标准恒星演化金属产额为 Y ，考虑到 AGB 星及超新星抛射出来的金属元素，再结合初始质量函数，他们假设每形成单位质量的恒星，会有比例为 R 的质量返回到冷气体中，同时产生比例为 Y 的金属元素。产生的金属元素立刻与周围的冷气体均匀混合。在前面所介绍的气体冷却、加热、抛射、再结合等过程中，金属元素都跟随气体流动，并且立即与周围气体混合。在这个过程中， Y 和 R 都是自由参数。

3 MCMC 方法在半解析模型中的应用

第 2 章列举了半解析模型涉及的一些重子物理过程，这些过程并不全面，但都是最基本的物理过程，一般在半解析模型中都会用到。还有一些物理过程，如盘的不稳定性，尘埃散

射等^[28, 29], 也会在不同的模型中用到。因为对这些物理过程的研究不够充分, 只能用简单的经验公式去描述, 这就导致了半解析模型中有众多的自由参数, 比如上一章提到的恒星形成效率 α , 类星体模式中的 f_{BH} , 超新星反馈中的 ϵ_{halo} , ϵ_{disk} 和 γ_{ej} , 活动星系核反馈中的 κ_{AGN} , 化学演化中的 Y 和 R 等。有些半解析模型^[30] 采用更复杂的公式描述这些重子物理过程, 会导致出现更多的自由参数。

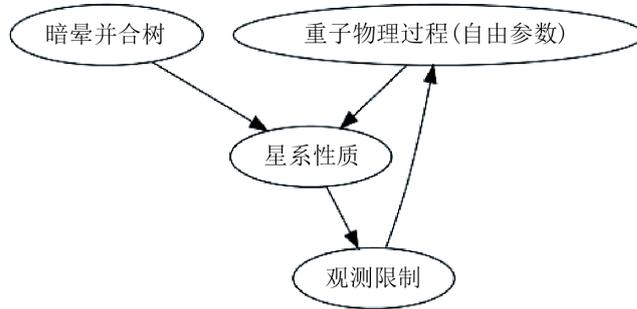


图 1 半解析模型示意图

半解析模型的一般过程如图 1 所示。模型输入是暗晕并合树^[31-33] 和重子物理过程, 在并合树最开始红移处, 每个暗晕中都有一定质量热气体, 之后热气体冷却形成星系; 根据暗晕性质和重子物理过程, 星系进行演化, 进而得到模型输出即星系的各种性质如质量、光度等。根据这些输出可以得到星系的大尺度统计性质, 如恒星质量函数、光度函数、相关函数等。将这些统计性质与观测结果进行对比, 如果有不相符的地方, 返回去调节模型中涉及物理过程的自由参数, 直至模型预测与观测相符。手动调节参数只能根据经验用眼睛对比结果, 这种方法费时费力; 而且有时调整某个参数使得模型预测的星系性质 (如恒星质量函数) 与观测相符, 但这可能导致另一个预测性质 (如相关函数) 与观测不符。这种参数间的相关性导致调整自由参数以使模型符合多种观测结果变得很困难。如果引入 MCMC 方法自动搜索参数空间, 不仅能大大提高半解析模型效率和准确度, 还能更好地理解这些参数是如何影响星系性质的^[34, 35]。通过 MCMC 方法可以得到参数的置信区间以及从统计学上确定模型预测与观测的相符程度^[36, 37]。MCMC 方法在天文学其他领域很早就得到了应用, 比如宇宙学参数的估算^[38], 也有不少研究者开发出成熟易用的软件包^[39-42]。

3.1 MCMC 方法

MCMC 方法即用于统计推理中进行模拟抽样, 它是 Markov Chain 和 Mento Carlo 两个概念的组合。Markov Chain (即马尔科夫链) 是一种随机过程, 它的一个重要属性就是无记忆性, 其所表示的随机过程在一个状态空间里游走且未来的状态只与当前的状态有关, 而与之前的状态无关, 这种无记忆性便称之为马尔科夫性。Mento Carlo 方法 (即蒙特卡洛方法) 可以视为一种思想的泛称, 只要在解决问题的过程中, 利用大量的随机样本, 然后对这些样本结果进行概率分析, 从而得到问题求解的方法, 都可以称之为蒙特卡洛方法。

MCMC 方法在贝叶斯推理中有着广泛的应用^[44, 45], 如算法模型的后验参数估计问题。

给定一组观测数据 D 和一组模型参数 x , 它的似然函数 $P(D|x)$ 表示在这组参数 x 下得到观测数据 D 的概率。贝叶斯理论认为, 给定一个先验概率 $P(x)$ 即参数 x 的先验分布, 后验概率 $P(x|D) \propto P(D|x)P(x)$, 就是说在观测数据 D 下得到参数 x 的概率正比于参数 x 的先验概率与似然函数的乘积。在半解析模型中, 重子物理过程中的每个参数都有一个范围, 这样我们就知道了参数 x 的先验分布。观测数据 D 即是观测得到的诸如恒星质量函数、光度函数等的星系性质, 从模型中我们也可以得到模型预测的这些星系性质, 两者结合可以得到似然函数 $P(D|x)$, 这样我们就可以得到后验概率。但是一个关键的问题是, 在取样的过程中, 如何从一组参数 x_1 跳到另一组参数 x_2 ? MCMC 方法正是解决这个问题的关键。

MCMC 方法用马尔科夫链来构造参数空间里参数点的跳转。假设后验概率 $P(x|D)$ 为 $\pi(x)$, 马尔科夫链构造了一个转移概率矩阵 Q , $q(i, j)$ 则表示从参数点 i 跳转到参数点 j 的概率。这个转移概率矩阵应该使马尔科夫链满足细致平稳条件:

$$\pi(x_i)q(i, j) = \pi(x_j)q(j, i) \quad . \quad (17)$$

显然, 通常情况下这个条件并不成立。我们可以引入一个 $\alpha(i, j)$, 使得:

$$\pi(x_i)q(i, j)\alpha(i, j) = \pi(x_j)q(j, i)\alpha(j, i) \quad . \quad (18)$$

按照对称性, 当我们取

$$\alpha(i, j) = \pi(x_j)q(j, i), \alpha(j, i) = \pi(x_i)q(i, j) \quad (19)$$

时, 式 (18) 就成立了。 $\alpha(i, j)$ 称为接受率, 可以理解为在原来的马尔科夫链上, 从参数点 i 以 $q(i, j)$ 的概率跳转到参数点 j 的时候, 我们以 $\alpha(i, j)$ 的概率接受这个转移。但是有一个小问题: 马尔科夫链在转移的过程中的接受率 $\alpha(i, j)$ 可能偏小, 使得采样过程中马尔科夫链容易原地踏步, 拒绝大量的跳转, 导致马尔科夫链遍历所有的状态空间需要花费太长的时间。解决这个问题需要同时扩大式 (18) 两边的接受率 α , 两边同时除以 $\alpha(i, j)$ 使得一边的接受率放大到 1, 这样我们就提高了采样中的跳转接受率。所以我们可以取

$$\alpha(i, j) = \min \left\{ \frac{\pi(x_j)q(j, i)}{\pi(x_i)q(i, j)}, 1 \right\} \quad . \quad (20)$$

这种算法就是著名的 Metropolis-Hastings 算法^[46], Kampakoglou 等人^[14]和 Henriques 等人^[47, 48] 在各自的半解析模型中用的就是这种算法。另一个著名的算法是 Gibbs Sampling^[49], 这里不做过多介绍, 有兴趣者可以参看相关文献。

3.2 半解析模型中使用 MCMC 方法

首次将 MCMC 方法应用到半解析模型中的是 Kampakoglou 等人^[14], 他们对 Daigne 等人^[50] 的半解析模型做了一些扩展, 总共是 7 个参数, 涉及到超新星反馈、恒星形成等物理过程。他们发现尘埃改正机制对于恒星形成过程模型的选取 (hierarchical 模型或者 monolithic 模型) 有着非常重要的作用。高效率的超新星外流可以使 hierarchical 模型更好地与观测相符, 而 monolithic 模型则需要贫金属星风的外流, 甚至是超大质量黑洞主导的外流。虽然

monolithic 模型的 χ_2 值 (描述理论预测与观测数据差距的量) 比 hierarchical 模型的大, 但是仍然不能将 monolithic 模型排除。

Henriques 等人^[47]于 2009 年将 MCMC 方法应用到 Munich 模型中。他们使用了 K 波段光度函数, $B - V$ 颜色分布和黑洞质量-核球质量关系这三个观测数据来限制模型参数。他们讨论了半解析模型的参数如何影响相应的星系特征, 并且发现这些参数之间有强烈的关联。他们发现, 要想使模型结果同时符合三个观测特征, 那么就需要比以前模型更高的超新星反馈效率。为了平衡, 他们又降低了低质量星系的反馈效率, 但是这又导致光度函数在 L_* 附近超过了观测值。为了能同时符合多种观测特征, 他们也讨论了 Munich 模型可能做的一些修改。之后, Mutch 等人^[51]也使用 MCMC 方法分析研究 Munich 模型, 他们则使用了 $z = 0$ 处质量函数, $z = 0.8$ 处质量函数和黑洞质量-核球质量关系作为观测限制, 他们发现, 只有把超新星反馈提升到大的不真实的值才能使模型与三个观测同时相符。据此他们得出结论, 认为现在的模型并不完整并且对模型的改正提出了建议。考虑了 Mutch 等人的建议, Henriques 等人^[48]修改了描述再结合过程的模型, 并用 MCMC 方法得到了同时符合不同红移多种星系观测特征的参数最佳拟合值。

Lu 等人^[34]也建立了一套应用 MCMC 方法的半解析模型。他们的模型包括 13 个自由参数, 并用当前红移的恒星质量函数作为观测限制。他们发现一些重要参数的后验概率分布像拓扑一样复杂并且简并, 说明要想探索参数空间需要对星系形成过程有足够的了解。他们还发现因为模型参数间的简并, 选择一个比较窄的先验分布能够给模型足够强的限制。他们通过模拟质量函数里的误差发现, 一个准确的观测误差模型对于作出有意义的推断至关重要。

下面我们以 Mutch 等人^[51]应用 MCMC 方法的 Munich 模型为例, 介绍一些成果。

3.2.1 单独观测限制

Mutch 等人的模型以 Croton 等人^[10]的模型为基础, 他们采用了 6 个自由参数, 在之前的章节中我们都提到过, 如表 1 所示。因为 Mutch 等人的模型使用的暗晕并合树来自于 Millennium 数值模拟, 运行一次半解析模型需要的时间很长; 而要想使 MCMC 链达到平稳状态, 需要运行模型的次数很多, 达到了近万次甚至几万次。因此使用整个模拟盒子是不现实的。Millennium 的输出分为 512 个文件, 每个文件包含的数据都是相互独立的, 因此 Mutch 等人随机选取一个文件作为模型输入, 这就大大降低了消耗时间。用了三个观测函数限制: $z = 0$ 处的恒星质量函数, $z = 0.83$ 处的恒星质量函数和 $z = 0$ 处的黑洞质量-核球质量关系。他们首先分别用这三个观测来限制参数范围, 最后共同使用三个观测找到参数的置信区间。

图 2 展示了只使用 $z = 0$ 处的恒星质量函数时 6 个自由参数的边缘后验分布。从图中可以看出, α_{SF} , ϵ_{halo} 和 ϵ_{disk} 的一维后验分布是一个类高斯分布; 这种情况说明这 3 个参数能够被 $z = 0$ 处恒星质量函数很好地限制。而其他 3 个参数并不是类高斯分布, 而是相对平坦的分布, 说明 $z = 0$ 处恒星质量函数并不能很好地限制 κ_{AGN} , f_{BH} 和 γ_{ej} 。图中的黄色圆点表示每个参数的最佳拟合值, 绿色十字表示 Croton 等人的模型中的取值。每个二维后验分布图里的两条黑色等高线则是代表 1σ 和 2σ 范围。 f_{BH} 和 κ_{AGN} 两个参数是描述两个黑洞增长

表 1 通过 MCMC 方法得到的半解析模型中 6 个自由参数的最佳拟合值以及置信区间

参数	Croton 等人的取值	$z = 0$ 处的最佳拟合值	$z = 0.83$ 处的最佳拟合值	$z = 0$ 加上 $z = 0.83$ 处的最佳拟合值
α_{SF}	0.07	$0.019^{+0.003}_{-0.003}$	$0.044^{+0.033}_{-0.019}$	$0.055^{+0.011}_{-0.016}$
k_{AGN}	5.89×10^{-4}	$1.90^{+0.39}_{-0.33} \times 10^{-4}$	$1.71^{+4.13}_{-1.21} \times 10^{-4}$	$1.47^{+0.24}_{-0.67} \times 10^{-4}$
f_{BH}	0.03	$0.015^{+0.002}_{-0.003}$	$0.015^{+0.035}_{-0.010}$	$0.025^{+0.007}_{-0.007}$
ϵ_{disk}	3.5	$5.14^{+1.22}_{-0.76}$	$4.79^{+1.87}_{-0.94}$	$13.8^{+4.1}_{-2.2}$
ϵ_{halo}	0.35	$0.26^{+0.06}_{-0.03}$	$0.41^{+0.16}_{-0.06}$	$1.18^{+0.38}_{-0.20}$
γ_{ej}	0.5	$7.1^{+4.9}_{-4.8} \times 10^{-3}$	$7.1^{+1.0}_{-4.2} \times 10^{-3}$	$1.13^{+0.30}_{-0.24}$

模式的参数，两者具有一定的相关性，虽然 $z = 0$ 处质量函数无法确定这两个参数的分布区间，但是对这两者的相关性做了很好的限制，如图 2 中 f_{BH} 和 κ_{AGN} 的二维后验分布函数所示，这两者具有一定的简并性。

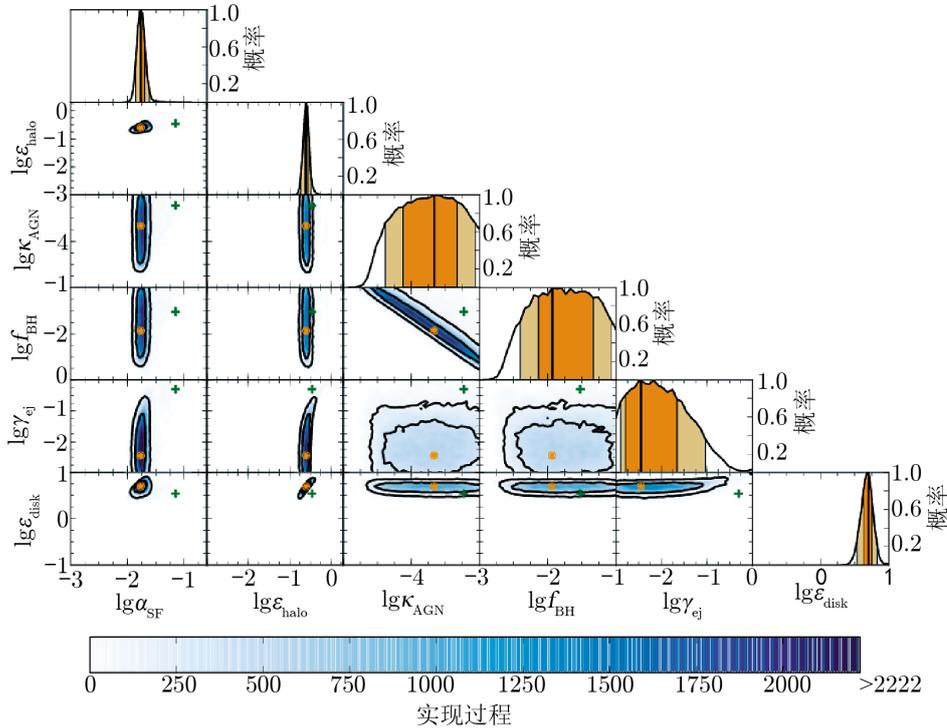
图 2 以 $z = 0$ 处恒星质量函数作为观测限制时，6 个参数的边缘后验分布^[51]

图 3 则显示了 $z = 0$ 处质量函数。带有误差棒的蓝色方块是观测数据，带有黄色区域的黑色实线是取最佳拟合值时的模型预测及 0.95 置信范围，红色虚线是 Croton 等人的结果，

黑色虚线则是参数取只用 $z = 0.83$ 处质量函数作为限制时的最佳拟合值时模型在 $z = 0$ 处的质量函数预测。很明显, 当只使用 $z = 0$ 处质量函数作为限制时, 最佳拟合值得到的结果与观测符合得最好。

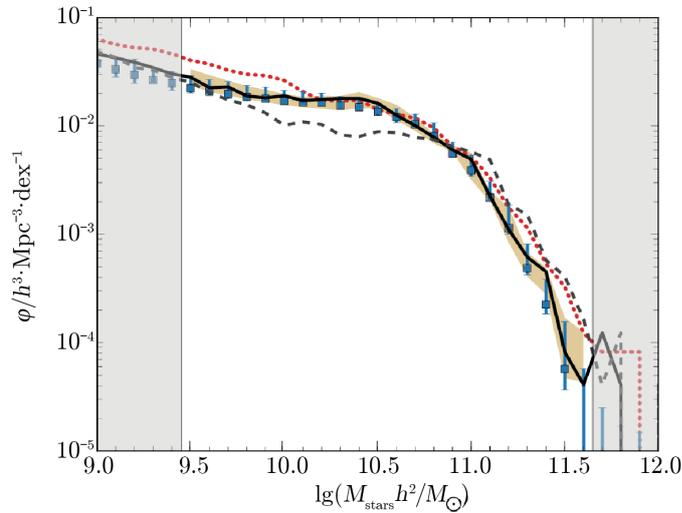


图 3 $z = 0$ 处恒星质量函数^[51]

为了打破 f_{BH} 和 κ_{AGN} 之间的简并, 需要一个能直接影响黑洞的观测数据来限制这两个参数。图 4 显示了只使用黑洞质量-核球质量关系作为限制时, α_{SF} 和 f_{BH} 的后验分布。图中两条黑色实线分别代表 α_{SF} 和 f_{BH} 的后验分布的 1σ 和 2σ 置信范围, 灰色虚线代表只使用 $z = 0$ 处质量函数作为限制时两者的后验分布置信区间, 红色实线代表当共同使用两个观测数据作为限制时两个参数的后验分布区间。不像 $z = 0$ 处恒星质量函数能对很多参数有所限制, 黑洞质量-核球质量关系只能限制 f_{BH} 和 α_{SF} 这 2 个参数。核球质量的增长有两种方式: 一种是通过并合, 但是模型中的 6 个自由参数没有一个是直接影响并合过程的; 另一种是通过星系盘的不稳定性。在 Mutch 等人的模型中, 他们设定当星系盘的恒星面密度超过一定值时, 盘就会不稳定, 盘中的一部分恒星就会进入核球, 因此核球质量可以通过这种增长方式与恒星形成效率 α_{SF} 联系起来。而黑洞质量的类星体增长模式主要受参数 f_{BH} 的影响。从图 4 可以看出, 当提高恒星形成效率 α_{SF} 时, 核球质量会通过盘的不稳定性增加, 则黑洞增长效率 f_{BH} 应相应增加。另外还能看出, 如果观测限制增多, 能对 α_{SF} 和 f_{BH} 这两个自由参数的置信范围做出更好的选择。可以相信, 如果把三个观测数据共同进行限制, 那么对这些自由参数的取值范围将更加确定。

3.2.2 联合观测限制

图 5 显示了当将 $z = 0$ 处质量函数, $z = 0.83$ 处质量函数和 $z = 0$ 黑洞质量-核球质量关系联合起来共同作为限制时, 6 个参数的边缘后验分布。正如所期待的那样, 6 个参数的一维边缘后验分布都有着类高斯的分布, 说明这三个观测数据对这 6 个自由参数做出了很强的限制。另外, 这 6 个参数之间的二维后验分布的置信范围也大大缩小。图 6 显示了当将三个

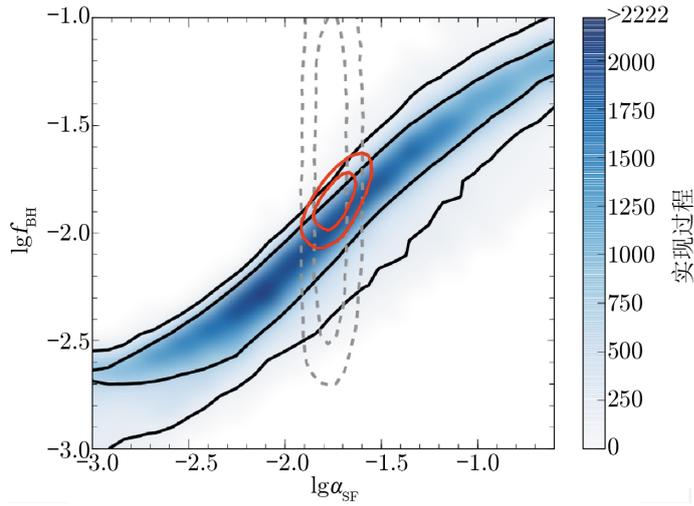


图 4 以 $z = 0$ 处黑洞质量-核球质量关系为限制时, f_{BH} 和 α_{SF} 这两个参数的后验分布^[51]

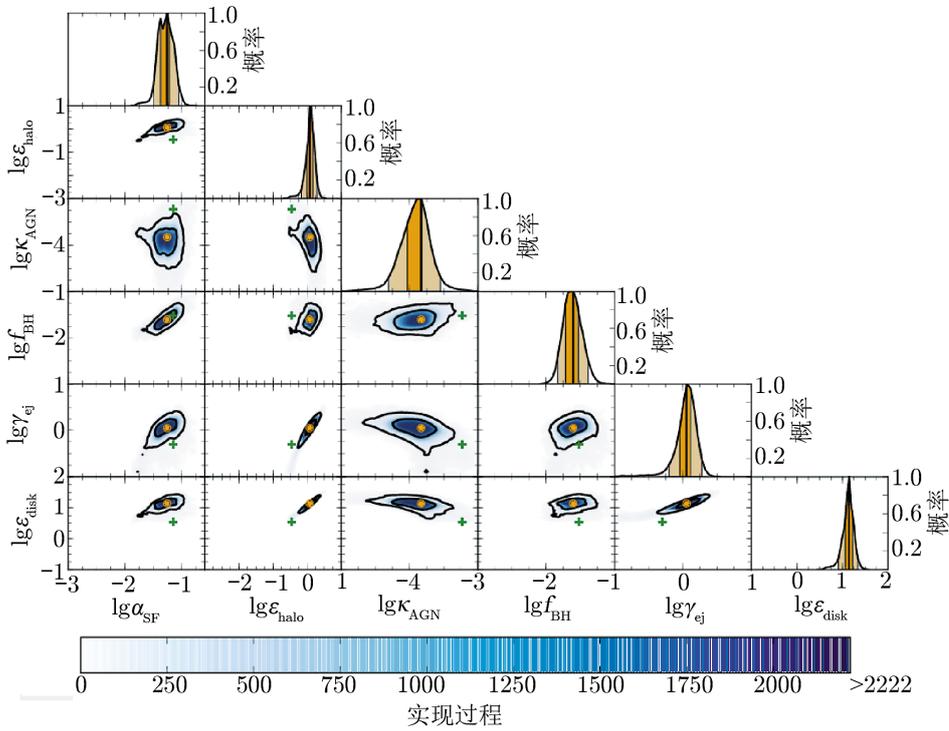


图 5 结合三个观测数据共同作为限制时, 6 个参数的边缘后验分布^[51]

观测数据结合共同作为限制时, $z = 0$ 处的质量函数。带有黄色区域的黑色实线是取最佳拟合值时的模型预测及 0.95 置信范围, 蓝色方块和红色虚线的意义与图 3 相同。可以看出, 模型预测与观测结果符合得很好, 只是在小质量星系处有稍微的超出。但是有一个问题需要注意: 虽然由 MCMC 方法得到了最佳拟合值及其置信范围, 但是有些参数的值看上去并不真实, 也就是说几乎没有物理意义, 这些最佳拟合值的大小参考表 1。比如 γ_{ej} 和 Croton 等人的取值相比, 当只用一种观测限制时小了将近 100 倍, 当使用三种观测限制时又大于 1。 γ_{ej} 代表暗晕的单位动力学时间内被再结合的气体的质量比, Croton 等人取 0.5, 代表在单位动力学时间内有一半的被抛出气体会再回到暗晕中。如果取大于 1 的数, 代表所有被抛出气体全部回到暗晕还不够。虽然 γ_{ej} 的取值没有物理意义, 但是也为进一步解决问题提供了思路: 如果要使 γ_{ej} 的取值有意义, 但又要同时满足不同的观测限制, 那么喷射出去的气体必须更快地回到暗晕中。Henriques 等人^[48] 对 Munich 模型中的再结合过程做了改进, 他们也应用了 MCMC 方法, 但是将式 (13) 中再结合的时标做了重新定义, 将之与红移联系起来, 即再结合的时标与红移有关, 引入了另外的参数, 使得 γ_{ej} 的数值大小变得有物理意义。

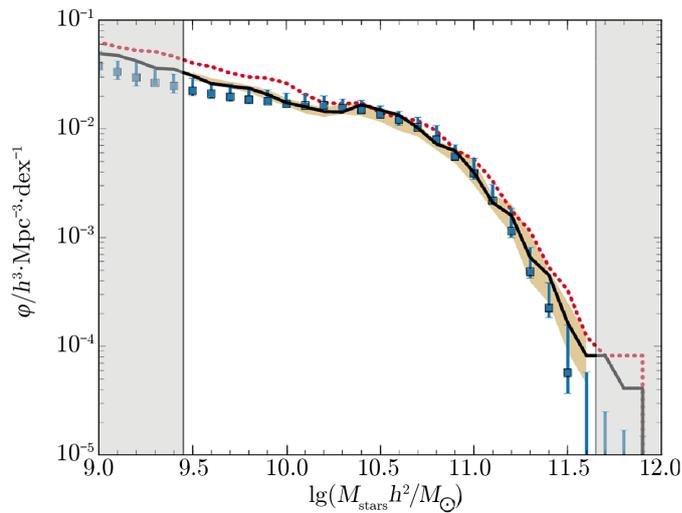


图 6 结合三个观测数据共同作为限制时, $z = 0$ 处的质量函数^[51]

4 总 结

本文介绍了 MCMC 方法在星系形成半解析模型中的应用。星系形成的半解析模型通过星系中的重子物理过程, 把暗物质晕和星系成功地连接起来, 对人们理解星系的形成和演化起到了很大作用。因为人们对重子物理过程的了解太少, 只能用经验公式去描述这些过程, 导致半解析模型中涉及到的自由参数太多。调节自由参数以使模型符合观测是半解析模型的通常做法, 但是过多的参数以及参数间的简并, 导致依靠经验手动调节参数变得很麻

烦。MCMC 方法提供了一种不用靠手而是自动搜索参数空间的路径, 把它应用到半解析模型中, 能大大提高模型的效率和确定参数的置信范围, 并且能从统计学上判定模型预测与观测的符合程度。限制模型的观测数据越多, 自由参数的置信区间就越好确定。总之, MCMC 方法能促进人们更好地使用半解析模型这个工具去理解星系的形成与演化。

参考文献:

- [1] Mo H, van den Bosch F C, White S. Galaxy formation and evolution. Cambridge: Cambridge University Press, 2010
- [2] Baugh C M. Reports on Progress in Physics, 2006, 69: 3101
- [3] White S D M, Frenk C S. ApJ, 1991, 379: 52
- [4] Kauffmann G, White S D M, Guiderdoni B. MNRAS, 1993, 264: 201
- [5] Mo H J, Mao S, White S D M. MNRAS, 1998, 295: 319
- [6] Somerville R S, Primack J R. MNRAS, 1999, 310: 1087
- [7] Cole S, Lacey C G, Baugh C M, et al. MNRAS, 2000, 319: 168
- [8] De Lucia G, Kauffmann G, Springel V, et al. MNRAS, 2004, 348: 333
- [9] Kang X, Jing Y P, Mo H J, et al. ApJ, 2005, 631: 21
- [10] Croton D J, Springel V, White S D M, et al. MNRAS, 2006, 365: 11
- [11] Bower R G, Benson A J, Malbon R, et al. MNRAS, 2006, 370: 645
- [12] Monaco P, Fontanot F, Taffoni G. MNRAS, 2007, 375: 1189
- [13] Guo Q, White S, Boylan-Kolchin M, et al. MNRAS, 2011, 413: 101
- [14] Kampakoglou M, Trotta R, Silk J. MNRAS, 2008, 384: 1414
- [15] Metropolis N, Ulam S. Journal of the American Statistical Association, 1949, 44(247): 335
- [16] Metropolis N, Rosenbluth A W, Rosenbluth M N, et al. J. Chem. Phys., 1953, 21: 1087
- [17] Somerville R S, Hopkins P F, Cox T J, et al. MNRAS, 2008, 391: 481
- [18] Nagashima M, Yahagi H, Enoki M, et al. ApJ, 2005, 634: 26
- [19] Kauffmann G, Colberg J M, Diaferio A, et al. MNRAS, 1999, 303: 188
- [20] Kauffmann G. MNRAS, 1996, 281: 475
- [21] Binney J, Tremaine S. Galactic dynamics. Princeton: Princeton University Press, 1987
- [22] Boylan-Kolchin M, Ma C P, Quataert E. MNRAS, 2008, 383: 93
- [23] Somerville R S, Primack J R, Faber S M. MNRAS, 2001, 320: 504
- [24] Cox T J, Jonsson P, Somerville R S, et al. MNRAS, 2008, 384: 386
- [25] De Lucia G, Kauffmann G, White S D M. MNRAS, 2004, 349: 1101
- [26] Fu J, Guo Q, Kauffmann G, et al. MNRAS, 2010, 409: 515
- [27] Fu J, Kauffmann G, Huang M L, et al. MNRAS, 2013, 434: 1531
- [28] Somerville R S, Gilmore R C, Primack J R, et al. MNRAS, 2012, 423: 1992
- [29] Baugh C M, Croton D J, Gaztañaga E, et al. MNRAS, 2004, 351: L44
- [30] Benson A J. New Ast., 2012, 17: 175
- [31] Springel V, White S D M, Tormen G, et al. MNRAS, 2001, 328: 726
- [32] Springel V. MNRAS, 2005, 364: 1105
- [33] Springel V, White S D M, Jenkins A, et al. Nature, 2005, 435: 629
- [34] Lu Y, Mo H J, Weinberg M D, et al. MNRAS, 2011, 416: 1949
- [35] Lu Y, Mo H J, Katz N, et al. MNRAS, 2012, 421: 1779
- [36] Benson A J, Bower R. MNRAS, 2010, 405: 1573
- [37] Bower R G, Vernon I, Goldstein M, et al. MNRAS, 2010, 407: 2017
- [38] Lewis A, Bridle S. Phys. Rev. D, 2002, 66(10): 103511

- [39] Lewis A. Phys. Rev. D, 2013, 87: 103529
- [40] Akeret J, Seehars S, Amara A, et al. Astronomy and Computing, 2013, 2: 27
- [41] Foreman-Mackey D, Hogg D W, Lang D, et al. Publications of the Astronomical Society of the Pacific, 2013, 125(925): 306
- [42] Kilbinger M, Wraith D, Robert C P, et al. MNRAS, 2010, 405: 2381
- [43] Robert C, Casella G. Statist. Sci., 2011, 26(1): 102
- [44] Ruiz de Austri R, Trotta R, Roszkowski L. Journal of High Energy Physics, 2006, 5: 002
- [45] Trotta R. Contemporary Physics, 2008, 49: 71
- [46] Hastings W K. Biometrika, 1970, 57(1): 97
- [47] Henriques B M B, Thomas P A, Oliver S, et al. MNRAS, 2009, 396: 535
- [48] Henriques B M B, White S D M, Thomas P A, et al. MNRAS, 2013, 431: 3373
- [49] Geman S, Geman D. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, PAMI-6(6): 721
- [50] Daigne F, Olive K A, Silk J, et al. ApJ, 2006, 647: 773
- [51] Mutch S J, Poole G B, Croton D J. MNRAS, 2013, 428: 2001

The MCMC Implementation in Semi-Analytical Model of Galaxy Formation

LI Shi-jie^{1,2}, YANG Xiao-hu³

(1. Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Shanghai 200030, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. Center for Astronomy and Astrophysics, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Semi-analytical model (SAM) of galaxy formation is an important tool to understand the baryon physics processes in galaxy formation and evolution. It forms and evolves galaxies in the dark matter halos using dark matter halo merger trees and baryon physics processes. Because of the poor understanding of baryon physics processes in galaxy formation, there are many free parameters in SAM. From SAM, people can get many galaxy properties such as stellar mass function, luminosity function, correlation function and so on. The model is usually adjusted by eyes to get the appropriate prediction which could agree well with the observation. This makes SAM very difficult and unreliable in determining model parameters.

MCMC (Markov chain Monte Carlo) method is an algorithm to run numerical simulation by random numbers which are generated by Markov chain. It is a Bayesian statistical technique for probing complex, highly degenerated probability distributions. It could sample a multidimensional space with a probability proportional to the likelihood that the model describes the observational constraints. The MCMC implementation in SAM could help to

search the free parameter space automatically and get more convinced parameter estimations. In recent years, many SAM researchers have applied MCMC method into SAM and got some promising results. These studies show that the MCMC method is an excellent tool to make SAM much more powerful in understanding the galaxy baryon physics processes.

Key words: galaxy formation; Semi-analytical model; MCMC method