

Understanding with respect to MC & MD

04/20/20

1) **Hierarchical paradigm of multiscale modeling** ---- Picture Source

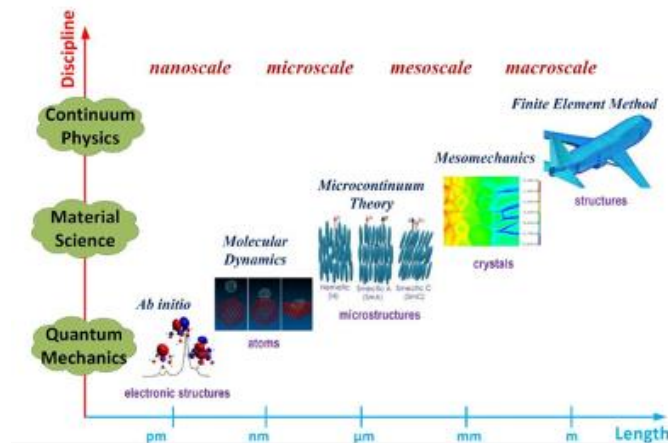
(a) <https://www.fhwa.dot.gov/publications/research/ear/13103/001.cfm>

(b) Original author ---- James Lee  
George Washington University

Lee, James D., Jiaoyan Li, Zhen Zhang, and Leyu Wang. "Sequential and concurrent multiscale modeling of multiphysics: from atoms to continuum." In *Micromechanics and Nanomechanics of Composite Solids*, pp. 1-38. Springer, Cham, 2018.

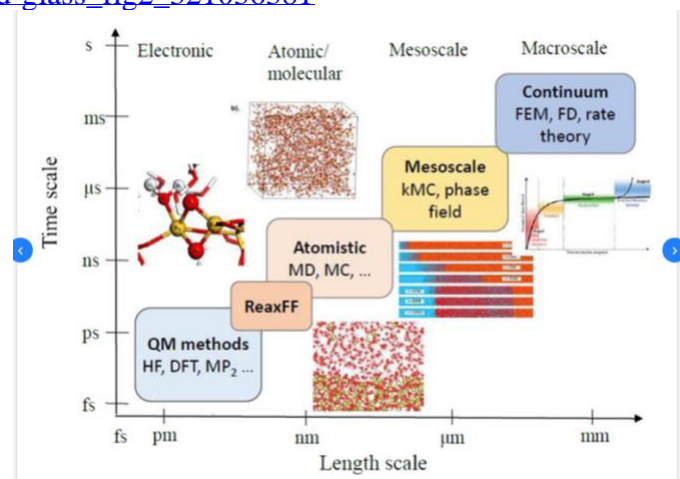
计算物理/计算化学中每处理一个体系之前，我们都要问我们想算什么，然后根据我们需要计算的性质来选择合适的工具，这样才能保证用有限的计算资源获得我们想知道的性质。

现在一个比较热门的词叫做“多尺度模拟” (multiscale modeling)，套用这张图 (图片出处：[xqwang.engr.uga.edu](http://xqwang.engr.uga.edu))



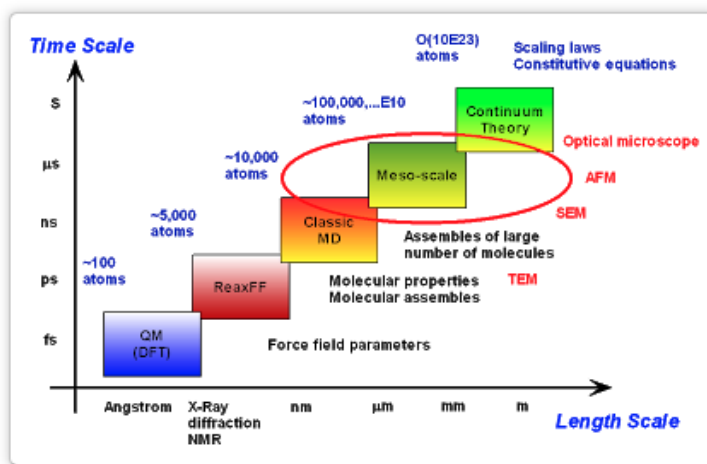
2) Another picture

[https://www.researchgate.net/figure/Multiscale-simulation-methods-and-applications-to-water-glass-interactions-and-glass\\_fig2\\_321036381](https://www.researchgate.net/figure/Multiscale-simulation-methods-and-applications-to-water-glass-interactions-and-glass_fig2_321036381)



### 3) Another Picture

<https://sites.google.com/site/yanglumultiscale/research-topics/multiscale-computational-mechanics>



Hierachical paradigm of multiscale modeling

2) (Metropolis 蒙特卡罗方法、动力学蒙特卡罗方法、分子动力学方法这三种模拟方法有何特点与差异?) (<https://www.zhihu.com/question/24334476>)

### Answer 1



刘书乐

编辑推荐

335 人赞同了该回答

这个问题很大，详细展开足够在一本分子模拟的教材里写满几章了，了解它们最好的方法应该是去啃书，比如分子模拟的经典教科书："Understanding Molecular Simulation, From Algorithms to Application" by Frenkel & Smit。

这里先讨论MD和Metropolis MC, 暂不讨论Kinetic Monte Carlo (KMC)。这是因为传统的MD和MC更具可比性（为简单起见，这里MC就是指狭义上的使用Metropolis算法的蒙特卡罗，而这里谈到的MD仅限于平衡态的MD模拟，暂不讨论非平衡态的NEMD），两种方法产生的初衷都是为了计算统计力学里的系综平均：假设系综所对应的相空间的概率分布为  $P(p, q)$  ( $p$ 为动量,  $q$ 为坐标)，那么对于任意平衡态热力学量  $M$ , 它的系综平均  $\langle M \rangle$  可以表示为

$$\langle M \rangle \equiv \int dp dq P(p, q) M(p, q) \quad (1)$$

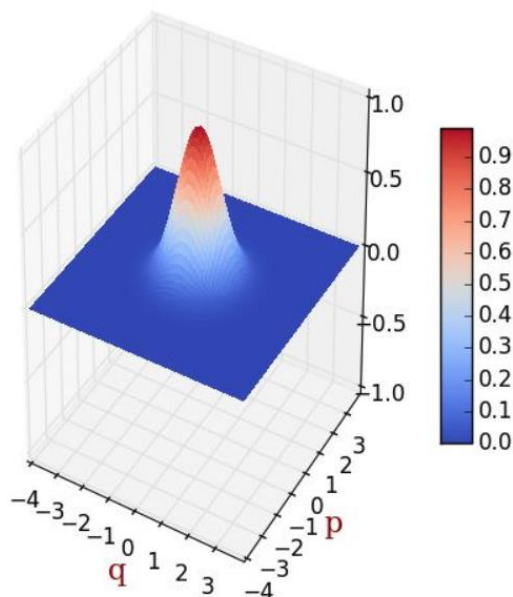
说得更直白一点，传统的MD和MC基本上是在干同一件事：算积分。只不过最早的MD是在微正则系综（NVE系综）里算积分，而最早的MC是在正则系综（NVT系综）里算积分。

先说MC。众所周知，正则系综里相空间的概率分布就是玻尔兹曼分布

$$P(p, q) = \frac{e^{-\beta \mathcal{H}(p, q)}}{\int dp dq e^{-\beta \mathcal{H}(p, q)}} = \frac{e^{-\beta \mathcal{H}(p, q)}}{Z} \quad (2)$$

其中  $\mathcal{H}$  是体系的哈密顿量， $Z \equiv \int dp dq e^{-\beta \mathcal{H}(p, q)}$  是配分函数， $\beta \equiv 1/k_B T$ 。以一维线性谐振子

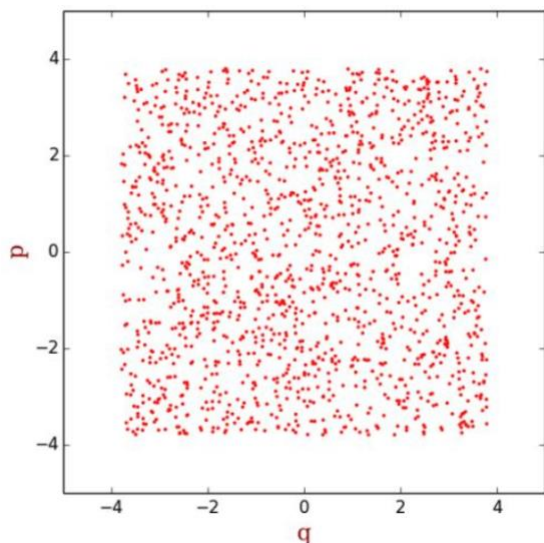
$\mathcal{H}(p, q) = \frac{1}{2} k q^2 + \frac{p^2}{2m}$  为例，这个相空间概率分布就是下图所示，其实就是个高斯分布。



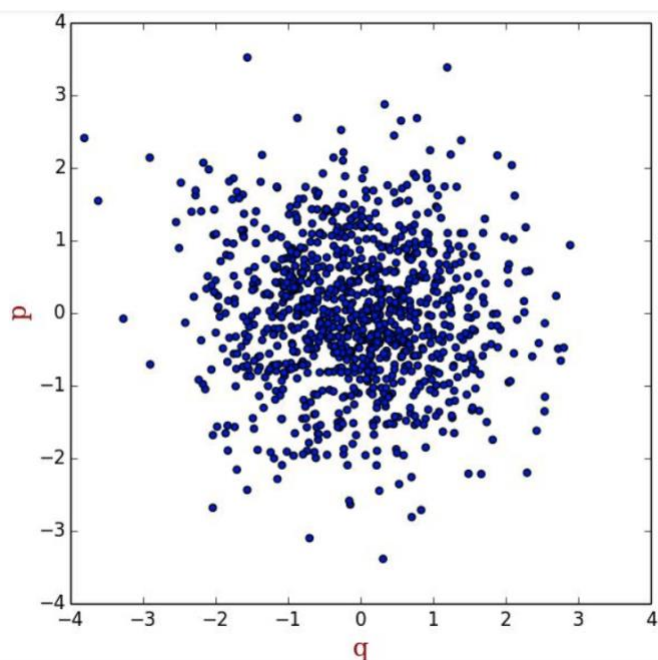
知道了概率分布，那么正则系综中任意平衡态热力学量  $M$  的系综平均  $\langle M \rangle$  就是下面这个积分

$$\langle M \rangle \equiv \int dp dq P(p, q) M(p, q) = \frac{\int dp dq M(p, q) e^{-\beta \mathcal{H}(p, q)}}{Z} \quad (3)$$

对于 (3) 式中的积分，绝大多数情况下是没有解析解的，因此只能通过数值求解。最简单粗暴的用蒙特卡洛方法求解 (3) 式积分的方法就是在整个积分区域上均匀地随机打点，求得每一点处的  $M(p, q) e^{-\beta \mathcal{H}(p, q)}$  值，然后对所有点求和再除以配分函数即可。例如对于一维线性谐振子，我们可以如下在相空间上均匀地打点



这种算法，的确非常简单，但效率很低。因为位于相空间中  $e^{-\beta\mathcal{H}}$  的值非常小的区域，系统处于这个区域的概率非常低，也就是说这些区域大体上来说对系统平均  $\langle M \rangle$  的贡献是非常小的，然而在采样过程中，由于我们是随机均匀打点，这样  $e^{-\beta\mathcal{H}}$  的值非常小的点和  $e^{-\beta\mathcal{H}}$  的值非常大的点的权重是一样的。换句话说，我们浪费了不少时间在那些权重很小的点上。那么是否有这样算法，能够使产生的点的分布符合玻尔兹曼分布，也就是说，在  $e^{-\beta\mathcal{H}}$  的值非常小的区域少打点，在  $e^{-\beta\mathcal{H}}$  的值非常大的区域多打点（如下图所示），然后把每一点处的  $M$  值相加，就能得到我们需要的系统平均呢？Metropolis算法就是为了实现这样的目标产生的。关于Metropolis算法的细节，我不在这里展开，很多书籍里都有描述。



再说MD。MD是基于牛顿方程的，而从牛顿方程是可以很自然导出孤立体系在保守力作用下是能量守恒的，因此最初的MD是在微正则系综(NVE)里计算系综平均。根据统计力学的等概率假设，对于总能量为E的微正则系综，其相空间概率密度分布为

$$P(p, q) = \frac{1}{\Omega(E)} \quad (4)$$

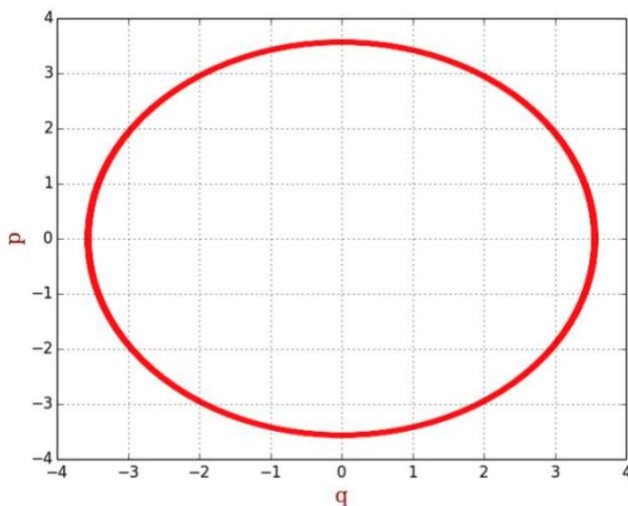
其中  $\Omega(E)$  是体系处于能量为E时的态密度(density of state)

$$\Omega(E) = \int dp dq \delta(E - \mathcal{H}(p, q)) \quad (5)$$

那么热力学量M在微正则系综里的系综平均  $\langle M \rangle$  就是

$$\langle M \rangle \equiv \int dp dq P(p, q) M(p, q) = \frac{1}{\Omega(E)} \int_V dp dq M(p, q) \quad (6)$$

其中V是积分的区域，也就是当能量为E时系统所占据的相空间。拿一维线性谐振子作为例子，当能量为E时其在相空间中区域对应于下图中的圆环，那么这个系综平均就是根据(6)式在对应于能量为E的圆环区域上进行积分。MD所做的，就是让粒子跑遍这个圆环上的每一个点(历经各态, ergodicity)，把每一点上M的值加起来求平均，得到的就是系综平均(time average equals ensemble average, this is what we call ergodicity)。



综上，MD和MC都可看作是为计算系综平均的采样方法。采用Metropolis算法的MC不局限于在正则系综里的采样，可以应用于其他系综，比如NPT和巨正则系综。至于微正则系综，理论上也可以使用Metropolis算法，只是由于各个态之间能量相等，概率也相等，这时使用Metropolis算法不能起到加速采样的作用，等于白用。同样地，MD的采样也不局限于微正则系综，当把一个孤立体系和一个恒温器(thermostat)或者恒压器(barostat)耦合起来，就可以进行正则系综或者NPT系综里的模拟。当然，如果耦合的恒温器或者恒压器是以extended Lagrangian形式实现的话(比如Nose-Hoover Thermostat)，广义上来讲MD模拟还是可看作在一个大的微正则系综中进行。



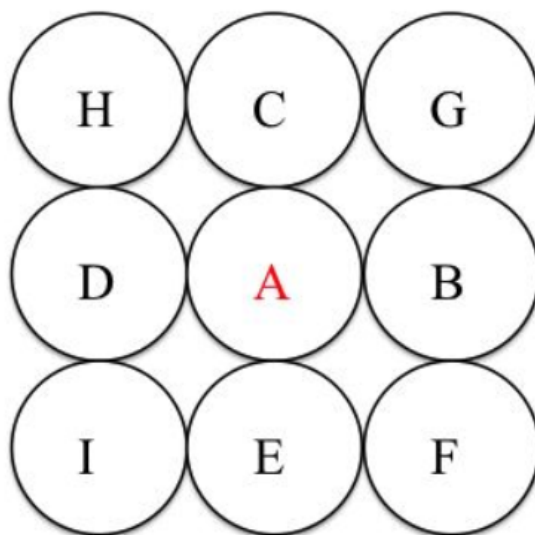
## MD和MC的比较

1. 算法的难易程度。一般说来，MD的算法比MC要复杂一些。这是因为，对于MC而言，每产生一帧新的构型只要计算能量然后根据acceptance ratio移动粒子就足够了。但对于MD，每产生一帧新的构型所需要的不只是能量，还需要每个粒子的速度和受力来对牛顿运动方程进行数值积分，这就加大了计算量。如果还要考虑为了提高计算效率的Verlet neighbor list等设置，MD的算法会更复杂一些。
2. 所能计算性质的多少。很明显，MD能比MC计算更多的性质，MC只能计算一些与动力学无关的平衡态性质，比如能量、比热容等等，MD既可以计算与动力学无关的性质，也可以计算与动力学相关的性质，比如扩散系数、热导率、黏度、态与态之间的跃迁速率，光谱等等。这是因为，使用Metropolis算法的MC，为了快速得到平衡态的分布，把整个系统简化为一条马尔可夫链，第N个构型只和第N-1个构型有关联，与第N-2或者更早的构型没有什么关联。在实际情况中不是这样的，系统是有记忆的，在 $t_0$ 时刻处于相空间中 $(p_0, q_0)$ 点的系统与在 $t_1$ 时刻处于 $(p_1, q_1)$ 的系统是有关联的，它们之间的关联可以用条件概率 $P(p_1, q_1; t_1 | p_0, q_0; t_0)$ 来描述。很多动力学性质都和这样的条件概率有关系，而MD产生的数据是可以计算这些概率分布的。所以世界上没有免费的午餐，MC虽然算法简单，但大体上只能计算静态性质，如果要计算动力学性质还是要上MD。
3. 可使用的软件。~~可使用的MD的软件远多于MC的软件。~~我个人觉得很大程度上是由于MD能够计算的性质远比MC多，并且数据可视化后给人更多直观的感受。比如说，你可以把MD模拟蛋白质折叠的轨迹文件用软件可视化之后做成动画，动画就比较直观地演示了蛋白质结构随时间的变化。可是如果你用MC做同样一个体系的模拟，你把轨迹可视化之后看到的可能就是一堆原子像无头苍蝇一样乱动，最后稳定在某个构型附近。一般来讲，MD软件的开发也不难，一旦把积分牛顿方程的engine写好，往里面添加新的相互作用势并不难，这样每添加一点新的功能，MD软件就可以适用于更多的体系。相比起MD软件的百花齐放，分子模拟中学术界通用的MC软件并不多，很多课题组都是自己写MC代码，但适用范围仅限于自己研究的体系。
4. 增强抽样(enhanced sampling)。这个问题 @JimKarrey 在他的回答中已经提到了。这是分子模拟中的一个坑，不管MD还是MC都会面临在系综中采样不足的问题。这方面知乎上讨论已经很多，可以参考下面的链接，在此不做详细描述。

计算化学领域中有哪些技术可以被称为是当前的黑科技？ - 知乎用户的回答  
兔子走遍世界——物理学视角下的采样方法 - 生命的设计原则 - 知乎专栏

最后说一下KMC，我在这方面不是很懂，期盼有高人来解答。

Kinetic Monte Carlo(KMC)和主要模拟平衡态性质的MD和Metropolis MC相差就比较远了，它主要模拟系统随时间的演化而非平衡态性质。KMC的基本思想是把整个系统划分成若干个态(state)，并获得态与态之间单位时间的跃迁几率（这个可以通过分子模拟或者过渡态理论来计算），然后通过具体的算法决定每一步往哪一个态跃迁，就可以模拟系统在各个态之间随时间演化。我没有做过KMC的研究，并不太了解KMC的算法，仅仅粗略看过一两篇用KMC模拟扩散的文章。以下的例子是我粗浅的理解，可能不太准确。



我们想用KMC模拟粒子在二维晶格上的扩散，并求出扩散系数。假定在单位时间步长内，粒子只能跳到它最相邻的格点。如图所示，粒子某时刻处于A格点，那么粒子下一时刻只有可能跳到B,C,D,E格点，单位时间内粒子从A跳到这些格点的跃迁几率分别为 $r_{AB}, r_{AC}, r_{AD}, r_{AE}$ 。结合这些跃迁几率，KMC的算法就可以通过生成随机数来决定粒子下一步所在的格点。在模拟足够长时间后我们可以算出粒子的均方位移(mean square displacement, MSD)，均方位移和扩散系数成正比关系，因而可以求出扩散系数。

当粒子只能在一维情况下运动（即只能从A到B或者从A到D），并且 $r_{AB} = r_{AD}$ ，这种情况下KMC其实就是在模拟一维随机游走(random walk)。

KMC的具体应用应该还很多，比如晶体生长什么的。



董玉龙

物理学 话题的优秀回答者

编辑推荐

yang元祐等 182 人赞同了该回答

这个问题非常好！我之前也想过一段时间，下面就把我的理解说一下。

还是先说一下为什么我觉得这个问题好。

1. 在利用统计物理的方法求系综平均时，肯定需要采样，而采样方法很多，到底应该用哪一种，这些方法的联系在哪，是否都是一种方法在不同条件下的应用（实际上都是MCMC）。
2. 上学期学了一些统计学习，里面也用到了很多采样方法，实际上，如果能把统计物理里的采样方法的理解提升到更高层次，完全是可以理解很多机器学习的问题的。传统机器学习里的Kalman滤波实际上就是用了Langevin动力学（这里的分子动力学也是一样的），Particle滤波用到的就是在重要采样的基础上发展起来的DMC。特别火的AlphaGo基于的RL，实际上就是对不同形式的Bellman方程进行采样，对价值函数进行估计，然后采用不同的策略，比如贪心算法， $\epsilon$ -贪心算法。怎么估计价值函数、做出决策，又发展了一系列的，比如Deep Q Network, Monte Carlo Tree Search。统计物理里在样本较少情况下常用的bootstrap方法也在RL中有体现，比如experience reply，这些都是AlphaGo算法里很核心的部分。

要理解这三个算法的不同，需要对统计里的采样方法有一点了解。一般的采样方法有

1. 直接采样：把累计分布函数的反函数拿来采样，但是C.D.F.一般不太容易得到，所以不常用。

$X \sim p(\cdot)$ , C.D.F.  $F(x) = \int_{-\infty}^x p(x)dx$ , 通过均匀分布采样  $x = F^{-1}(u)$ ,  $u \sim U(0, 1)$

2. 接受-拒绝采样：拿一个大的函数把待采样的分布套起来（因为有归一化要求，必须要用一个大于1的数乘上某个分布） $G(x) = c \cdot g(x)$ , s.t.  $G(x) \geq p(x)$ ,  $\forall x$ , &  $\int g(x)dx = 1$ 。采样按照拒绝率判断是否接受采样样本。采样位置  $x \sim g(\cdot)$ ，接下来判断是否接受， $u \sim U(0, 1)$ ，如果  $u \leq \frac{p(x)}{G(x)}$ ，接受x，否则，拒绝x。相对于直接采样法效果好一些，但是，要找到好的G(x)不是简单的事情。如果c太小了，很难满足要求，如果c太大了，接受率太低，采样效率不高

3. 重要采样：通过变换  $\mathbb{E}_p[f(X)] = \int f(x)p(x)dx = \int f(x)\frac{p(x)}{g(x)}g(x)dx = \mathbb{E}_g[f(X)\frac{p(X)}{g(X)}]$ ，  
通过按照g(x)采样，对带权的函数求平均来求之前p(x)下的期望。重要采样是为了解决p(x)采样的样本点比较密集处f(x)的值较小导致的偏差较大的情况。但也有不足，就是g(x)不好找，并且，在某些情况下，采样次数要比较多才能收敛。

4. MCMC的方法：这里的三种方法都属于MCMC，Markov Chain Monte Carlo。都是要构造一条Markov链，使其稳定分布收敛到待采样的分布，这样，收敛之后每一步的结果都是从目标分布p(x)中采样的样本。Markov链能收敛的必要条件就是细致平衡条件成立，即  
 $p(x, t|y, t-1)p(y, t-1) = p(y, t|x, t-1)p(x, t-1)$



下面就分别说下题主提到的三种采样方法。

用到的一些关于微观动力学方程的知识可以在专栏里找到。

董玉龙：还还没想好标题3--非平衡统计物理1

zhuanlan.zhihu.com

知

## 一、Metropolis

1. Metropolis-Hasting算法：Metropolis算法实际上是Metropolis-Hasting算法的特殊情况。MH算法的核心在于，我们通过提议分布来控制采样位置，接受率来控制是否进行采样，跟接受-拒绝采样比较像。假设有便于采样的提议分布  $g(y|x)$ 。在当前采样时刻t，需要知道上一步t-1的样本点x，t时刻的样本点  $y \sim g(\cdot|x)$ 。均匀分布随机数  $u \sim U(0, 1)$  与接受率

$\alpha(x, y) = \min\{\frac{g(x|y)p(y)}{g(y|x)p(x)}, 1\}$  进行比较，如果  $u \leq \alpha(x, y)$ ，接受y，否则，拒绝y。

2. Metropolis算法：当提议分布是对称的情况，i.e.  $g(y|x) = g(x|y)$ ，有  $\alpha(x, y) = \min\{\frac{p(y)}{p(x)}, 1\}$
3. 物理中的Metropolis算法：由于在物理里，我们的提议分布都采用高斯分布等对称分布的形式。所以，MH算法退化到Metropolis算法。
4. 应用：用Metropolis算法计算Ising模型，每次是否接受自旋反转，都是通过稳定分布（正则分布）之比来验证。

## 二、KMC

1. 为什么要引入KMC，Kinetic Monte Carlo。回忆一下之前说的MH算法相当于接受-拒绝采样，而接受-拒绝采样的不足之处在于接受率会不高，比如当模拟退火时，温度太低，可能会导致接受率非常低。
2. KMC的优势：实际上KMC相当于Metropolis算法在时间上进行粗粒化。每次不判断是否接受提议，而是通过轮盘赌的方法选出一个一定接受的操作，并且计算这个操作所需要的时间。相当于，在Ising模型的例子中，有可能1000步才会接受一次自旋反转，但是在KMC里不跑1000步，1步接受，并且估计这一次操作所需要的时间为1000步。
3. 应用：用的比较多的是在化学反应系统中。系统中有  $\{R_i, r_i\}_{i=1}^N$  个反应通道，每个R对应一个化学反应方程，r是该反应的速率，根据浓度和速率常数计算。轮盘赌抽样

$u \sim U(0, 1)$ ,  $\frac{\sum_{i=1}^{k-1} r_i}{\sum_i r_i} \leq u < \frac{\sum_{i=1}^k r_i}{\sum_i r_i}$ ,  $k = 1, 2, \dots, N$  返回满足条件的k，则这一步，反应

$R_k$  发生，其对应的分子相应变化，这一步的时间， $u \sim U(0, 1)$ ,  $\Delta t = \frac{\ln(1/u)}{\sum_i r_i}$

### 三、分子动力学, Molecular Dynamics, MD

1. 首先要解释一下，题主在提问里说，MD是确定性算法，这是错误的。首先想一下MD怎么工作的，是不是用到了高斯随机数。其次，MD会被误认为是确定性算法，还是由于Langevin方程非常类似于牛顿方程，这一方面是一件好事，即是你不懂Ito公式，不懂随机过程，也可以通过牛顿力学来理解MD，创造了不少科研岗位，但另一方面，门槛低了，有很多人误解做计算，特别是MD不需要什么数学物理知识，只是用用软件，调调参数。
2. Gibbs算法的引入：首先介绍另一个MCMC算法，Gibbs算法，不是Gibbs发明的，是在研究Gibbs随机场的情况下发明的。之前说过MH算法的不足在于有接受率，所以需要改进，引入了KMC，在时间尺度进行粗粒化，但是从上面的例子也可以看出，KMC能采样的系统是不够精细的，上面例子中你并不知道到底哪个分子在什么位置发生了化学反应，只知道系统的浓度发生了变化。因此，在应用中，KMC是不够的。另外，由于分子系统的自由度非常大，相当于要在6N维的相空间上进行采样，接受率即使是0.99，也会产生维数灾难，使得整个样本向量的采样接受率几乎为0。要解决接受率的问题，就要从接受率本身来看，如果能使提议分布满足 $p(y)g(x|y) = p(x)g(y|x)$ ，接受率就是1。Gibbs算法就是从这个角度出发。特别的，如果你的提议分布就是Markov链的转移概率，因为MCMC中的Markov链满足细致平衡条件，Gibbs算法就成立了。
3. Gibbs算法：记待采样的样本向量为  $x = (X_1, X_2, \dots, X_n) \in \mathbb{R}^n$ ，记t时刻采样时，除去第j个分量的向量为  $x_{-j}^{(t)} = (X_1^{(t)}, \dots, X_{j-1}^{(t)}, X_{j+1}^{(t-1)}, \dots, X_n^{(t-1)})$ ，采样过程中，在t时刻， $X_j^{(t)} \sim g(\cdot | x_{-j}^{(t)})$ ,  $j = 1, 2, \dots, n$ ， $x^{(t)} = (X_1^{(t)}, X_2^{(t)}, \dots, X_n^{(t)})$  为t时刻采样得到的样本向量。这个算法实际上就是在  $\mathbb{R}^n$  中沿各个坐标轴分别随机游走。
4. 应用：在MD中，提议分布采用高斯分布，因为从Langevin方程离散得到的向前差分采样就是从高斯分布里采样。稳定分布则通过Langevin方程对应的Fokker-Planck方程得到。这两个概率密度是满足细致平衡条件，也就是Gibbs算法中对提议分布的要求，所以，每次采样都接受。因此，在MD中每一步构型改变都是在相空间里进行了一次Gibbs采样。

## 四、补充点东西

1. 前面提到的几种基本采样方法，实际上在统计物理里也有很多应用，比如重要采样就衍生出了像DMC (Diffusion MC)等方法，在高维空间采样中有比较重要的应用。不过用DMC还是有很多要注意的技巧的，如果不注意，DMC的采样结果非常容易收敛到退化分布。
2. 三种方法的特点：Metropolis在Ising模型及其类似的体系里用的很多，好处是微观细节反映的比较好，但是因为接受率的问题，采样速度慢，并且只在低维空间比较有效。KMC在有平均场近似的地方用的很多，比如复杂系统、反应扩散系统、用统计物理研究社会问题。采样速度比较快，在要研究的时间尺度比较大时可以考虑，但是因为只体现了平均场近似之后的性质，对微观过程的体现不太好，而且也只适用于低维空间。MD对微观细节的体现非常好，适合研究像polymer、protein这样的生物物理模型，因为没有拒绝率的问题，在高维空间采样的效率和稳定性都比较好。并且，MD能最大化的得到系统的几乎所有力学量，因为这个采样方法几乎是在系综（统计物理里对概率空间  $(\Omega, \mathcal{F}, \mu)$  的描述）上给出了样本的足够多次实现（Markov链收敛之后，想要采样多少次都可以），根据强大数律，力学量的样本均值a.s.收敛到期望。
3. MD的基础上有很多发展，比如各种保辛算法的应用，通过Markov链轨道概率的分解发展出来的Path Integral MD，在时间、空间上的粗粒化方法。

被李老师 @Yongle Li 指出微正则系综里的分子动力学是确定性算法，i.e. 哈密顿力学。不过个人理解，实际上是Langevin方程的特殊形式。由于微正则系综下不考虑能量的涨落，所以扩散过程因扩散张量  $B(z, t) = 0$  而退化到确定性过程，而扩散张量也是噪声的关联张量，因为其为0，噪声也就退化为取值为0的退化分布。因此就不需要考虑噪声项。Langevin方程退化为  $\frac{dX(t)}{dt} = A(X(t), t)$ ，在确定性过程中，根据Liouville定理，概率密度对应的Fokker-Planck方程的解为退化分布，取值在前面常微分方程的解上。而前面的方程即为哈密顿正则方程

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases}$$

部分内容见专栏

董玉龙：非平衡统计1.5——从Markov到各种方程

zhuanlan.zhihu.com

知

### Answer 3



小侯飞氖

物理学、科研 话题的优秀回答者

编辑推荐

yang元祐等 75 人赞同了该回答

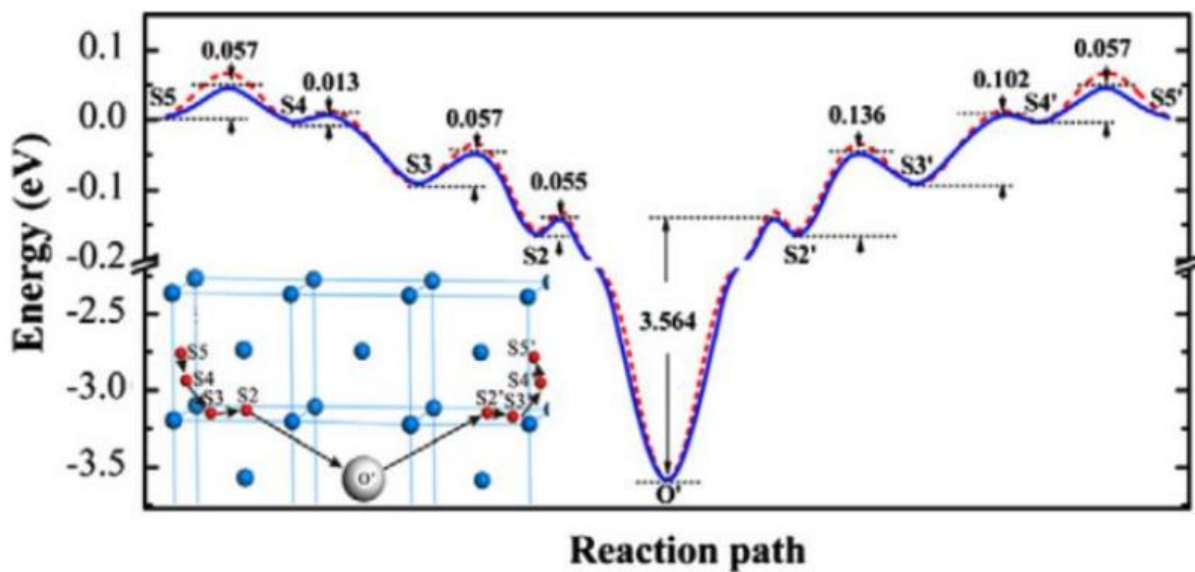
MD和MC了解不深，上面也有比较详细的回答了，我就不强答了。KMC这块我稍微了解一些，自己也写了几千行的KMC代码，说说自己的理解吧。

1、特点：

Kinetic Monte Carlo (KMC)是专门用于模拟体系长时间演化的一种MC方法，其核心思想和MC是一样的：利用伪随机数来模拟热扰动，只不过在KMC中能量极小值附近的热振动是被忽略掉的，KMC仅仅考虑引起组态跃迁（一个极小值到另一个极小值）的热扰动而已。

KMC有几个常见的假设：

a) 绝大多数情况下，体系原子都在能量极小值附近做准简谐的热振动，这样的能量极小态被称为体系的组态，例如：



Mo中He在空位附近的能量分布



b) 数学上，极小值之间必然有一个极大值或鞍点，所以从一个组态跃迁到另一个组态需要跨越一个能垒，跨越这个能垒的能量由热扰动提供。

c) 两次跃迁之间，原子只是在能量极小位置附近振动，这类振动是KMC不关心的，是被忽略掉的。

很显然，KMC适用于模拟“由组态跃迁主导的体系的长时间演化”，例如缺陷在晶体内部和表面的扩散、聚集、分解等。而类似于蛋白质折叠这种并没有明显组态变化的物理体系，就不适合用KMC来进行模拟。

## 2、算法简介：

一般情况下，KMC模拟中每个可能的跃迁能垒都要提前输入（有的算法会调用分子静力学等方法实时计算能垒）。在确定所有跃迁途径的能垒后，给出每个跃迁的反应速率。某个反应的反应速率由Arrhenius equation给出：

$$v^i = v_0 \exp(-E_b^i / k_b T)$$

其中  $E_b^i$  是第*i*个跃迁的能垒， $v_0$  为尝试频率，可以近似的取为反应物振动频率。然后根据速率随机选择跃迁途径，选中第*i*个跃迁途径的概率正比于其反应速率：

$$p(i) = v^i / \sum_j v^j$$

执行该跃迁反应，更新反应带来的跃迁途径及其能垒的变化，并按下式增加整个体系的模拟时间：

$$\Delta t = 1 / \sum_j v^j$$

以上为一个完整的Monte Carlo步，程序会重复以上过程，直至达到所需时间或者步数为止。

## 3、用途：

由于仅仅考虑引起组态跃迁的反应，KMC模拟能有效滤掉无效的热振动，大大的提高模拟效率。

以ising模型为例，传统MC算法会进行一次次反转自旋的尝试，然后根据metropolis判据选择是否接受该尝试。然而在低温下，这种尝试大部分都会被拒绝掉。假设拒绝的概率为999/1000，那么metropolis算法要进行1000次尝试才能得到一次有效反转，剩余999次都是无效的热振动。而KMC则相反，它不管三七二十一先反转了，然后根据反转概率（跃迁速率），给出此次反转所需尝试次数（1000次）。KMC用一步就完成了MC模拟1000步干的活，效率自然特别高。

因此，KMC能处理宏观时间尺度问题（秒~年）的模拟。

当然有利也有弊，由于体系中除去组态跃迁之外的所有信息都被忽略掉了，因此要求体系的演化轨迹必须是由这些已知的组态跃迁主导，而无法考虑未知的组态跃迁。举例来说，假设我们用KMC模拟单个空位V在晶体内的扩散系数，此时考虑的组态跃迁只有空位迁移一个。然而实际情况下，晶体中可能有位错、晶界等缺陷，这些缺陷能与空位相互作用影响其扩散系数，如果没有提前将该机制纳入考虑的话，模拟中是体现不出该影响的。因此，正确的设计模拟所关心的物理机制是KMC模拟的核心。

一言以蔽之，KMC一般不用来做机理研究，而是在已知跃迁物理机制的情况下，模拟多种跃迁机制对体系长时间演化的影响。它常与DFT、MD、MS等较为“精细”的方法结合，形成一个从原子到微米尺度，从飞秒到年尺度的多尺度模拟方法。细节可参考Nature Materials上的这篇文章：

#### Multiscale modelling of defect kinetics in irradiated iron

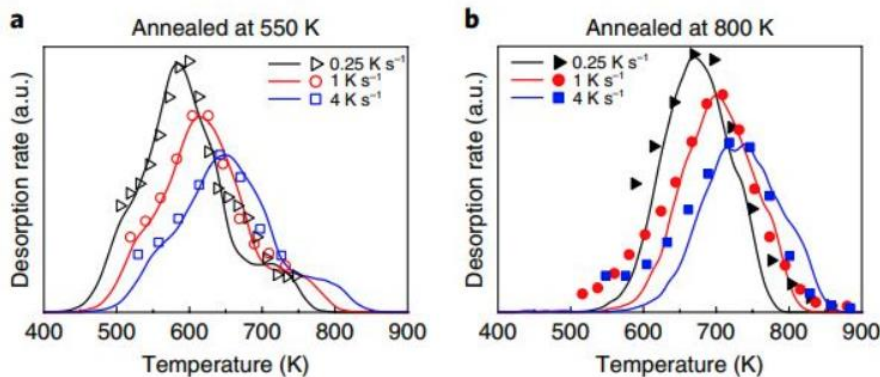
[www.nature.com](http://www.nature.com)



如上所述，KMC模拟中每个可能的跃迁能垒都要提前输入，这些能垒的准确性往往决定KMC模拟结果的有效性。

这篇文章研究氢从金属中辐照缺陷中的脱附行为，由于辐照产生的纳米孔洞通常是藏氢纳垢的罪魁祸首，氢和纳米孔洞的相互结合是该问题的物理核心。因此，我们先基于大量DFT计算，建立起了一个的氢-纳米孔洞结合物理模型。

用该模型可以快速准确的预测氢-孔洞结合能，从而给出氢的脱附能垒，输入到KMC当中去，结合宏观的实验条件（退火时间/温度/辐照剂量/升温速度等）进行大尺度模拟，从而给出与实验吻合的模拟结果：



**Fig. 4 | Thermal desorption spectra of deuterium from W. a,b**, Samples were irradiated at 10 keV per D ion to a fluence of  $3 \times 10^{19} \text{ D m}^{-2}$ , annealed at 550 K (**a**) and 800 K (**b**) for 5 min to render small and large nanovoids, implanted with 0.67 keV per D ions to a fluence of  $10^{19} \text{ D m}^{-2}$ , then annealed at different heating rates. Lines are multiscale modelling results and symbols are experimental data from refs. <sup>12</sup> (open) and <sup>13</sup> (filled).

## Answer 4



杨子微

北京大学 理学博士

从技术的角度来讲，想要真正理解这个问题，要明白MD（Molecular Dynamics分子动力学）和MC（Monte Carlo蒙特卡罗）的最大差别即可：MC只需要算势能（Potential Energy），MD除了算势能还需要算力。就是因为这一最大差异，导致MC和MD的应用范围不同，特点也不同。

因为可以算力就可以算速度，所以MD的可以由当前的坐标和速度计算出下一步的坐标和速度，也就是说MD可以得到时间连续的轨迹。

因为没有力也就没有速度，所以MC只能靠随机移动来决定下一步的位置，然后通过势能的不同计算每一步的权重，或者根据前后两步势能差的大小来决定是否接受这一随机移动的概率（Metropolis MC）。

从统计学上来讲，两种方法对体系的抽样是等价的。然而MD相比MC可以得到时间连续的轨迹，所以MD可以计算的性质更多。而且MD是根据力按部就班地一步一步计算，即使是对于复杂的真实体系也很容易实现MD软件的通用化。相比而言，因为复杂体系的随机移动是非常困难的（越复杂的体系在坐标随机移动之后就容易出现重叠，这一点是随机移动所不允许的），基本上不同的体系需要完全不同的随机移动算法，所以MC软件非常难以实现通用化。

然而MD必须算力这一点同样也制约了MD在某些领域的应用，因为力的计算需要势能对体系坐标是连续可导的，然而并不是所有体系的都可以满足这个条件。而MC则不需要算力，而且因为MC的实现远比MD简单，虽然没有通用的软件，针对特定体系编写一个MC程序往往也不复杂。所以MC特别适合自建模型体系的模拟，因为模型体系的势能有可能十分复杂，无法或者很难求导，而且模型体系的随机移动算法也更容易实现。

综上所述，MC和MD最根本差别就是是否算力。需要算力的MD程序更加复杂，但是可以实现通用化，且可以得到时间连续的轨迹，所以适合于模拟复杂的真实体系。而不需要算力的MC程序则比较简单，即使体系势能无法求导仍然可以模拟，但是随机移动算法难以实现通用性，导致MC程序也很难实现通用性，所以MC更适用于模型体系。

至于KMC（动力学蒙特卡罗）我本人并没有真正用过，所以无法回答的太深。然而KMC本质上仍然是MC，使用MC计算不同“态”之间的势垒从而间接获得“动力学”，然而却仍然无法得到时间连续的轨迹。



## Answer 5



Jim Karrey  
量子物理幼儿园小班。

做模拟的目的一般来讲，一个是理解、观察过程的微观机理，再一个是从微观状态统计某些过程的宏观量。

首先说分子动力学（MD）模拟，他的基本框架就是经典力学的运动方程积分的确定论的方法，但体系本身的演化是一个自然的马尔可夫链（多体过程就不要想着有确定的解析解了）。MD在恒温恒压控制的算法方面也会引入大量随机过程。常见的是Langevin dynamics（恒温）和Langevin piston（恒压），做MD模拟的看到这两个名词简直再熟悉不过。

做MD的模拟的目的是什么？首先是在微观层面上观察到分子运动，进而可以理解某些过程的机理，这一点是MD的巨大优势。其次，但也是最重要的，MD需要实现各态遍历性（Ergodicity）从而能够给出一些宏观统计量的值。每一个宏观统计量，像自由能，熵，焓，扩散系数，结合常数等等，都对应着某些微观量的系综统计平均。要得到一个准确的系综统计平均，就要实现各态遍历性。如何保证一个MD模拟可以实现各态遍历？这个问题直到二十世纪末才直到得到回答。一个MD方法应该具有辛性（Symplecticity，不要问我这个词什么意思，它就是音译）才能遍历各态，到这一步，对MD的基础理论的研究已经属于数学上的辛几何了，有志青年可以研究研究，个人感觉经典力学发展到这一步，一点都不比量子力学简单，数学形式上可以还要更加艰深。

解决了各态遍历的可实现性，现在的问题就是如何实现各态遍历性，或一定条件下的（伪）各态遍历。于是现在就有了多种加速采样方法。包括像Umbrella Sampling, Metadynamics, Adaptive Biasing Force之类的方法。个人已被ABF玩死。这些算法的确是“加速”了，但是想要得到一个好的统计结果，需要下花很多工夫，操作起来比直接的MD模拟麻烦多了。

MD是很好，结果准确（相对蒙卡而言），过程直观。缺点就是，计算资源消耗极大，做计算一定要找个土豪钱多的组，多买或者多租计算节点，要不然就会像我这样被坑惨了。

再说蒙卡方法，这是一类基于概率论的方法。我没有真正做过蒙卡的模拟，就凭我的了解说一点。讲错了轻拍。

大家都熟悉的一个蒙卡方法的例子是这样的：一张正方形白纸上画了一个圆，现在要求计算这个圆的面积。正方形的面积是容易计算的，很好办。圆就难多了。采用蒙卡的思想的话，可以往纸上随机地扔图钉。当图钉数量很多时，可以统计在圆内和圆外的图钉的比例，从而大致估计圆的面积占据整张正方形白纸面积的比例，于是就能估算得圆的面积了。这里，圆的面积就是我们要求得的统计量，而扔图钉则是我们产生的随机过程。大伙还可以用类似的方法写个程序计算圆周率的值。

蒙卡方法应用于分子模拟中，则是这样的：随机产生体系的状态，在每个状态（即每一次采样）中计算统计量的值。当采样数足够多的时候，计算这些统计量的Boltzmann系统平均，然后就得到了相应的宏观物理量的估计值。因此可以看到，蒙卡方法应用于分子模拟中，是直接从各态遍历性下手的。产生一堆状态，并不太管体系在这些状态里演化的每一步的细节。因此，蒙卡很难给出一个直观的体系在状态间演化的过程，而是需要你根据蒙卡的统计结果间接推测演化的路径和机理。

由于蒙卡是随机方法，所以每次估计出来的值都不会一样。但对于一个已经敛的采样过程，再增加大量的采样数目，最后的数值也只会在一个较小的合理区间内波动。

一般的蒙卡方法应用于简单的分子体系是很好办的。比如一个纯水的盒子这样的。随机产生水分子的分布就可以了。但是这种简单的蒙卡方法的局限性也是非常明显的，即它会产生大量的不合理的体系状态，这些状态在统计上的贡献极低，从而浪费了大量的计算资源。我们知道，一个物理量F的系统平均是这样计算的：

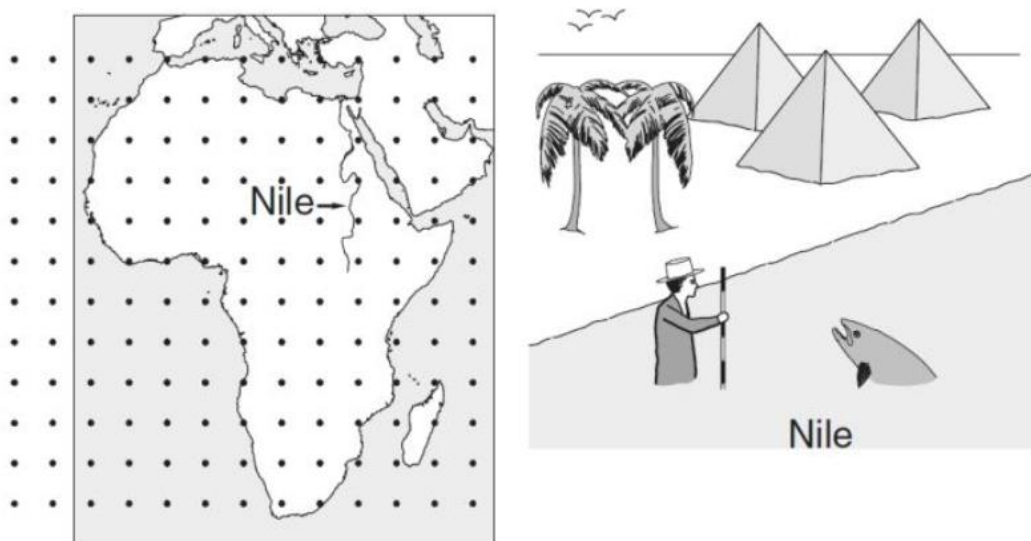


$$\langle F \rangle = \frac{\int F(p, q) \bullet e^{-\frac{E(p, q)}{k_B T}}}{\int e^{-\frac{E(p, q)}{k_B T}}}$$

能量越高的地方，权重越低。要知道，一个体系的自然演化路径，总是天然地要找能量最低路径。  
只有这些能量最低路径上的采样才是最有用，最有效的。于是就有了Metropolis蒙卡方法。

Metropolis蒙卡方法则是先得到体系的一个合理状态，然后尽可能地沿着一个合理的路径进行采样，从而大大提高采样的效率。

下图是我从一部经典的分子动力学教材中截取的插图（来自《Understanding Molecular Simulation, From Algorithms to Application》，作者Daan Frenkel, Berend Smit）：



**Figure 3.1:** Measuring the depth of the Nile: a comparison of conventional quadrature (left), with the Metropolis scheme (right).

想要测量尼罗河的平均深度，我们不必在整个地图上随机采点测量（左）。站在河里面，沿着合理的路径（河道）来采样才是最高效的。

简单粗暴的蒙卡方法对于生物蛋白质这样的体系来说是非常致命的。简单粗暴地随机产生构象的话，很容易就破坏了蛋白质的二级结构，而你希望采样到的正确的构象则往往无法恢复。这时候，如果考虑蛋白质本身的行为性质，循着它本身构象变化的路径采样，那样才会得到比较正确的结果。（这一段完全是我的推测，期待用蒙卡做过蛋白质模拟的人来细说。）

比较MD和MC，除了理论框架的确定论和概率论的差别以外，一个重要的区别是，MD天然地产生一个马尔可夫链，而MC的马尔可夫性是需要额外的方法来补充的。

## Answer 6



退乎  
已退乎

27 人赞同了该回答

这个问题很大，你应该去弄几本计算物理的书，要真正说清楚的话一个分别就是一个单元。

我分别用最简单的话给你说下那是在干啥。真正要学，找计算物理学的书自己去看。这类的书没有特别好读的。你说的这三种算法，每一个都是100多页的内容。你别听俺下面扯好像觉得自己懂了，每一个都是深坑。够你钻研几个月的。当然你判断一个人是不是真正懂一门学问，听他讲就行了。如果是没有任何基础的外行都能窥见其中的门径，那么不懂简直就是不可能的。这一点譬如费因曼先生讲的普通物理，水平好一点的中学生也是可以听的很明白晓畅的。

分子动力学算法：

首先你要明确，分子动力学算法是建立在经典力学模型上头的。是不考虑量子效应的。这个问题里外加的就是一个经典力学的力（势场）。

算法：

1. 起始状态（体系内粒子的一些初始物理信息，比如位置和速度），此时时间为 $t$ 。
2. 取一个极短的时间：外力（场）作用于粒子上，粒子发生运动。这一步也被称作蛙跳算法，leap-frog algorithm. 时间在这一步骤里的变化量是 $\Delta t$ 。
3. 粒子完成了运动，时间变成了 $t+\Delta t$ ，这个时候其物理信息（比如位置和速度）是可以计算的。
4. 你现在拥有所有点的物理信息，那么自然可以通过一些物理方法计算，整个体系的一些状态量比如内能。Okay, let us do it! 不用害怕，这种东西一般是在计算机上完成的，不需要你用手做。
5. 反复重复2-3的过程。在计算机编程上自然就是一个循环了。

★ 蒙特卡罗算法是什么：最简单的一句话就是，用生成随机数替代高维情况下的积分问题。比如，你弄一个日本的地图。他的边缘是非常复杂的曲线。如果对他进行强行积分的话很不好办。那么怎么办呢。就是用这种思路。把日本的图形放在一个比如二维平面里。之后再做生成随机数求面积这种运算。

★ 蒙特卡罗和分子动力说白了都是在算积分。但是区别在哪？分子动力是从时间的观点上考虑问题的，而蒙特卡罗是在能量的观点上考虑问题的。

那么接下来要搞清楚的自然就是Metropolis算法和动力学算法的区别了。这个稍微有点tricky！最准备数据库的考试，回头来填坑。

3) 分子动力学模拟的结果是确定的么？

<https://zhuanlan.zhihu.com/p/63076211>

## 2.2 分子动力学模拟的计算原理

人们对于大自然的认识程度，或者说物理学发展进步的衡量尺度可以是人们对于物理现象的观测尺度。这也是为什么理论物理学家在研究夸克等极小的粒子，而天体物理学家专注于观察星系的演变和宇宙的边缘。这样的尺度不仅仅存在于时间，也存在于空间当中。分子动力学模拟的计算元素是原子，与第一性原理考察电子场，求解薛定谔方程不同，在分子动力学模拟的世界里，原子就是组成这个世界的最小元素，不可再分。通过第一性原理计算或者实验，我们获得分子动力学模拟的力场，在基于牛顿经典力学，我们便可求解固定时间步原子与分子的物理运动，通过获得粒子的位移矢量和动量矢量，在统计力学的方法上求得所需要计算的物理量。简单来说，分子动力学是基于牛顿力学在热力学系统下研究原子与分子的物理运动的模拟方法。

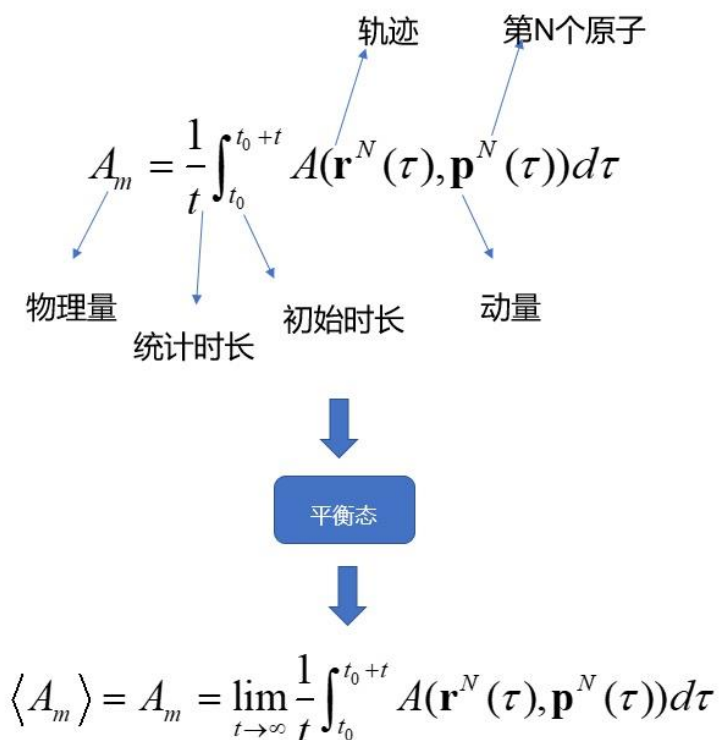


图 6 分子动力学模拟中物理量的统计