

摘要

本篇综述文围绕 Quantum Monte Carlo (QMC) 介绍了求解非相对论情形下 Schrödinger 方程的多个思路和方法.

非相对论量子力学的核心问题是 Schrödinger 方程的求解 [2]:

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = H|\psi\rangle.$$

式中 $|\psi\rangle$ 表示受考虑量子系统的态, 在通常的空间坐标下 (R 表象) 可以表示为波函数 $\langle \mathbf{x}|\psi\rangle = \psi(\mathbf{x})$. 这个方程只有极少数的情形存在解析解 (如自由粒子、简谐振子和氢原子等), 而大多数的情况都要以来数值计算 [1] 的方法.

Schrödinger 方程的变分形式

我们考虑定态 Schrödinger 方程的解, i.e. 本征值问题:

$$E|\psi\rangle = H|\psi\rangle. \quad (1)$$

我们可以使用通常用以求解 PDE 的方法如有限差分法 (FDM) 和有限元素法 (FEM) 去求解方程 (1), 但是这类方法在处理方程 (1) 时往往会遇到稀疏程度非常高的迭代矩阵 (往往这类矩阵是如此的稀疏以至于使用常用的 sparseBLAS 类方法也会造成大量的资源浪费), 不易处理. 因此, 变分方法在很多情形下是求解 Schrödinger 方程的更有效的手段.

量子态 $|\psi\rangle$ (或者说, 波函数 $\psi(\mathbf{x})$) 满足定态 Schrödinger 方程 (1) 等价于系统能量在这个态下的的数学期望

$$E[\psi] = \frac{\langle \psi|H|\psi\rangle}{\langle \psi|\psi\rangle} = \frac{\int \psi^*(\mathbf{x})H\psi(\mathbf{x})d^3x}{\int \psi^*(\mathbf{x})\psi(\mathbf{x})d^3x}. \quad (2)$$

在态 ψ 附近取稳定值, i.e. 若这个态改变了一微小的量 $\delta\psi$, E 期望的改变量不存在一阶项:

$$\delta E[\psi] \approx 0. \quad (3)$$

为了证明上面这一论断, 我们定义

$$P = \langle \psi|H|\psi\rangle, Q = \langle \psi|\psi\rangle,$$

将 δE 写到 $\delta\psi$ 的一阶项:

$$\begin{aligned} \delta E &= \frac{\langle \psi + \delta\psi|H|\psi + \delta\psi\rangle}{\langle \psi + \delta\psi|\psi + \delta\psi\rangle} - \frac{\langle \psi|H|\psi\rangle}{\langle \psi|\psi\rangle} \\ &\approx \frac{\langle \delta\psi|H|\psi\rangle - (P/Q)\langle \delta\psi|\psi\rangle}{Q} + \\ &\quad \frac{\langle \psi|H|\delta\psi\rangle - (P/Q)\langle \psi|\delta\psi\rangle}{Q}, \end{aligned} \quad (4)$$

结合 (2), (4) 和 $E = P/Q$, 我们可以得到:

$$H|\psi\rangle = E|\psi\rangle,$$

也就是定态 Schrödinger 方程 (1).

定态波函数位于 Hilbert 空间的一个子空间内. 如果用线性变分: 选取一组规范正交基 $|\chi_i\rangle, i = 1, \dots, N$ s.t.

$$\langle \chi_i | \chi_j \rangle = \delta_{ij}, \quad (5)$$

则一个态可以表示成如下的线性叠加 (注意这里我们使用了 Einstein 求和约定: 角标相同则对该角标求和, 下同):

$$|\psi\rangle = C_j |\chi_j\rangle. \quad (6)$$

那么能量泛函的值由下式给出 (式中 $H_{ij} = \langle \chi_i | H | \chi_j \rangle$):

$$E = \frac{\sum_{i,j} H_{ij} C_i^* C_j}{\sum_{i,j} \delta_{ij} C_i^* C_j}, \quad (7)$$

而式 (1) 化为如下形式:

$$\mathbf{HC} = \mathbf{EC}, \quad (8)$$

即式 (1) 的求解被划归为了特征值问题 (8). 求解这个问题所得到的每个特征向量 $\mathbf{C}_i^{(a)}$ 所对应的特征值 $E_i^{(a)}$ 满足

$$E_i^{(a)} \geq E'_i, \quad (9)$$

其中 E'_i 表示该系统第 i 个本征态所对应的真实能量 (方便表示起见, 这里所有的解均按能量 (本征值) 从低到高的顺序 $i = 0, 1, \dots$ 排列). 之所以我们所获得的解与真值有所不同, 是因为我们所取得基所张的空间只是 Hilbert 空间的一个子空间, 如果我们扩展现有的基, 使其成为 Hilbert 空间更大的一个子空间, 那新求得的能量本征值 $E_i^{(b)}$ 应该满足

$$E_i^{(a)} \geq E_i^{(b)} \geq E'_i.$$

注意到如果我们把基函数去成一个个的 Dirac Delta 函数, 实际上我们就得到了通常意义上的差分方程的计算格式, 但这种基在求解这类问题时非常低效, 其原因便是文章开头提到的超稀疏矩阵问题. 而通常用于进行变分问题计算的基都是经过仔细选取的连续函数. 求解特

征值问题的最优解法一般需要 $\mathcal{O}(n^3)$ 级的计算量.

另外, 如果我们所取的基不是完全正交归一的, 而是存在关系

$$\langle \chi_p | \chi_q \rangle =: S_{pq},$$

那么式 (8) 就要变形成为

$$\mathbf{HC} = \mathbf{ESC},$$

这种问题被称为广义特征值问题.

通常, 为了满足全同 Fermion 的 Pauli 不相容关系, 基函数会被选取成一系列反对称 (奇宇称) 波函数. 著名的 Hatree-Fock 方法就是基于这个思想实现的.

矩阵方法的精度问题

尽管变分方法把 Schrödinger 方程的求解问题简化成为了一个特征值问题 (或者广义特征值问题), 这种方法的计算复杂度非常的高: 当收到考虑的电子数目 n 上升时, 描述多体系态的波函数 $\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ 维数也会线性地上升, 而此时正确描述系统状态所需要的基的数目必须以指数形式上升. 综上, 当电子数目上升时, 矩阵方法求解变分形式的 Schrödinger 方程的计算量会变得非常大.

变分 QMC 方法

从这一节开始, 我们将放弃考虑矩阵形式的方程 (8), 转而在优化的思想讨论如何找到使得式 (2) 中的能量 E 取得最小值的那个态 $|\psi\rangle$. 在正式开始 Quantum Monte Carlo 方法的介绍之前, 我们先来类比一般的优化问题来简要构造一下基于变分方法求解 Schrödinger 方程的大致思路:

1. 基于参数集 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_S)$ 构造试探多体波函数 $\psi_{\boldsymbol{\alpha}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$.
2. 求能量的期望

$$E[\psi_{\boldsymbol{\alpha}}] = \frac{\langle \psi_{\boldsymbol{\alpha}} | H | \psi_{\boldsymbol{\alpha}} \rangle}{\langle \psi_{\boldsymbol{\alpha}} | \psi_{\boldsymbol{\alpha}} \rangle}.$$

3. 利用所得结果对 α 进行评估, 循环.

评估能量期望和更新参数集的方法这里暂时不予以讨论, 我们先来考虑上面步骤中求能量期望的这一步. 在这里, 我们需要计算式

$$E[\psi_\alpha] = \frac{\langle \psi_\alpha | H | \psi_\alpha \rangle}{\langle \psi_\alpha | \psi_\alpha \rangle} = \frac{\int \psi_\alpha^*(\mathbf{x}) H \psi_\alpha(\mathbf{x}) d^3x}{\int \psi_\alpha^*(\mathbf{x}) \psi_\alpha(\mathbf{x}) d^3x}. \quad (10)$$

中的积分. 这是一个高维积分, 采用传统的插值积分或者高维梯形积分的方法不仅构造繁琐, 在 n 上升时程序还不具有通用性. *Monte Carlo (MC)* 方法是一类利用随机数来实现某一类计算的方法. 在处理高维积分的时候, MC 积分具有精度随点数线性增长的特点; 而重要抽样方法中的 Metropolis Monte Carlo 方法可以通过细致平衡获得符合任意分布的随机数序列.

由于真实波函数在 Hilbert 空间的大部分区域的值都非常小, 我们在求解积分式 (10) 时又会不可避免地遇到稀疏问题. 此时, 如果采用传统积分方法或者直接使用均匀随机数进行 Monte Carlo 积分的话, 会造成大量的计算资源的浪费. 在这种情形下, 采用 Metropolis Monte Carlo (MMC) 方法进行积分是比较可取的: 使用大量粒子在波函数空间上进行随机游走, MMC 所对应的游走规则会将它们“推向”波函数值较大的区域, 达成较为高效的采样积分的目的.

我们定义如下局部能量:

$$E_{L\alpha}(\mathbf{x}) = \frac{H\psi_\alpha(\mathbf{x})}{\psi_\alpha(\mathbf{x})}, \quad (11)$$

注意如果 ψ_α 与系统真实的基态 $\psi(\mathbf{x})$ 相等, 这个局部能量是一个与 \mathbf{x} 无关的常数. 这样, 系统能量的期望就能被写成如下的形式:

$$E[\psi_\alpha] = \frac{\int \psi_\alpha^2(\mathbf{x}) E_{L\alpha}(\mathbf{x}) d^3x}{\int \psi_\alpha^2(\mathbf{x}) d^3x}. \quad (12)$$

有了式 (12) 形式的积分, 我们就可以利用分布

$$\rho(\mathbf{x}) = \frac{\psi_\alpha^2(\mathbf{x})}{\int \psi_\alpha^2(\mathbf{x}') d^3x'} \quad (13)$$

进行 MMC 抽样计算.

变分 QMC 方法直接利用基态能量最小原理计算多体系统的基态, 它的优点是直观而清晰, 但是显而易见的缺点是这个方法对构造和修正波函数的质量依赖较为严重.

投影 QMC 方法

除了基于评估和修改试探波函数的变分 QMC 方法, 我们还可以将 Schrödinger 的求解诉诸迭代法的思想, i.e. 通过一系列变换使得被研究系统的波函数 $\psi^{(i)}(\mathbf{x})$ 收敛到它的基态 $\psi'(\mathbf{x})$.

为了说明这一过程, 我们先考虑单个粒子的自由随机游走: 考虑一个位于一维离散时空间

$$\begin{aligned} x &= n\Delta l, \quad n = \dots, -1, 0, 1, 2, \dots; \\ t &= m\Delta t, \quad m = 0, 1, 2, \dots, \end{aligned}$$

中的粒子, 假设粒子向左或向右移动的概率都是 β , 而停留在原来位置的几率为 $1 - 2\beta$, 那么在 t 时刻找到粒子的概率服从 Markov 过程主方程:

$$\begin{aligned} p(x, t + \Delta t) - p(x, t) & \quad (14) \\ &= \beta[p(x - \Delta l, t) + p(x + \Delta l, t) - 2p(x, t)] \\ &\approx \beta\Delta x^2 \frac{\partial^2 p}{\partial x^2}. \end{aligned}$$

等式 (14) 最左边在 Δt 很小时也可以近似写成 $\Delta t(\partial p / \partial t)$, 再记 $\gamma = \Delta x^2 \beta / \Delta t$, 我们可以得到方程 (14) 的近似连续形式:

$$\frac{\partial p(x, t)}{\partial t} = \gamma \frac{\partial^2 p(x, t)}{\partial x^2}, \quad (15)$$

我们马上发现这是一个含时扩散方程. 于是我们可以就此写出它的 Green 函数:

$$G(x, x', t) = \frac{1}{\sqrt{4\pi\gamma t}} \exp\left(-\frac{(x - x')^2}{4\gamma t}\right), \quad (16)$$

使得

$$p(x, t) = \int G(x, x', t) p(x', 0) dx'.$$

在这里我们注意到, 我们可以通过如下方法构造另一个符合上述方程的 Markov 过程: 取粒

子 0 时刻的位置为 x_0 , 令它在 $+\Delta t$ 时刻游走到 x_1 的概率密度为 $G(x_1, x_0, \Delta t)$; 完成这部游走之后, 令它下一步 (时刻 x_2) 游走到 $+2\Delta t$ 的概率密度为 $G(x_2, x_1, \Delta t)$, 以此类推. 用符号表示这一 Markov 过程就是

$$T_{\Delta t}(x' \rightarrow x) = G(x, x', \Delta t). \quad (17)$$

该 Markov 过程和原先的随机游走的区别在于它利用了方程 (15) 的连续解的形式, 使得它具有理论上更高的效率. 这里我们再使用两个式子概括式 (17) 所描述的 Markov 过程:

$$x(t + \Delta t) = x(t) + \eta\sqrt{\Delta t}, \quad (18)$$

式中 η 是一个满足方差为 2γ Gaussian 分布随机变量:

$$p(\eta) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left(-\frac{\eta^2}{4\gamma}\right).$$

扩散方程的一般形式是:

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} = \mathcal{L}\rho(\mathbf{x}, t), \quad (19)$$

式中 \mathcal{L} 是一个二阶线性微分算子. 方程 (19) 具有如下的形式解:

$$\rho(\mathbf{x}, t) = \exp(t\mathcal{L})\rho(\mathbf{x}, 0). \quad (20)$$

当然, 由于形式解 (20) 中含有微分算符的指数项, 不能直接进行数值求解, 但是从这个方程我们可以发现, 方程 (19) 的 Green 函数就是式 (20) 中的微分算符 $\exp(t\mathcal{L})$ 在 R 表象下的矩阵元:

$$G(\mathbf{x}, \mathbf{x}', t) = \langle \mathbf{x}' | \exp(t\mathcal{L}) | \mathbf{x} \rangle. \quad (21)$$

我们自然会想到使用 Green 函数的某种近似去构造 Markov 过程求解扩散方程 (19), 但是注意, Green 函数只有可以被归一化, i.e. $\int G(\mathbf{x}, \mathbf{x}', t) d^3x = 1$ 时才可以被用来构造 Markov 过程. 我们会在下面对投影 QMC 的正式讨论中遇到这个问题.

投影 QMC 方法考虑如下的扩散方程:

$$\frac{\partial \rho(\mathbf{x}, \tau)}{\partial \tau} = \frac{1}{2} \nabla^2 \rho(\mathbf{x}, \tau) - V(\mathbf{x})\rho(\mathbf{x}, \tau), \quad (22)$$

注意到, 如果我们在方程 (22) 中令 $\tau = it$, 这个方程就成为了单位质量粒子的含时 Schrödinger 方程. 利用式 (20), 我们可以把解写成:

$$\rho(\mathbf{x}, \tau) = \exp[\tau(-K - V)]\rho(\mathbf{x}, 0), \quad (23)$$

其中 $K = p^2/2 = -1/2\nabla^2$ 是动能算符. 由于 K, V 不对易, 式 (23) 的指数项不能直接被展开, 但是, 如果我们令 τ 取一个很小的值, 那么指数项中的 Campbell-Baker-Hausdoff (CBH) 对易子就可以被忽略:

$$\begin{aligned} \exp[-\tau(K + V)] & \\ &= \exp(-\tau K) \exp(-\tau V) + \mathcal{O}(\tau^2). \end{aligned} \quad (24)$$

现在, 我们的主要任务就成为了找到式 (24) 右边各指数项的显式表达式. Green 函数势能项的指数展开是平凡的, 问题仍然出在如何找到动能项所对应的矩阵元:

$$G_{\text{kin}}(\mathbf{x}, \mathbf{x}', \tau) = \langle \mathbf{x}' | \exp(-\tau K) | \mathbf{x} \rangle. \quad (25)$$

为了做到这件事, 我们在式 (25) 中插入完备正交归一化关系:

$$\int |\mathbf{p}\rangle \langle \mathbf{p}| d^3p = 1,$$

并利用基本投影关系:

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{\sqrt{2\pi}} \exp(i\mathbf{p} \cdot \mathbf{x}),$$

由于动能算符 K 在 P 表象下是对角化的, 我们可以算得:

$$G_{\text{kin}}(\mathbf{x}, \mathbf{x}', \tau) = \frac{1}{\sqrt{2\pi\tau}} \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2\tau}\right). \quad (26)$$

我们发现, 式 (26) 作为虚时间 Schrödinger 方程 (22) 在不含势能项 V 情况下的 Green 函数实际上就是不含时扩散方程的解 (16), 这也使我们确信了它作为正确解的地位. 现在我们把虚时间 Schrödinger 方程 (22) 的 Green 函数再写一下:

$$\begin{aligned} G(\mathbf{x}, \mathbf{x}', \tau) & \\ &= G_{\text{kin}}(\mathbf{x}, \mathbf{x}', \tau) \exp(-\tau V) + \mathcal{O}(\tau^2). \end{aligned} \quad (27)$$

但不幸的是, 势能项因子 $\exp(-\tau V)$ 破坏了 Green 函数 (27) 的归一性. 对式 (27) 进行归一化的方法是在其上乘一个收敛因子 $\exp(\tau E_T)$ (但是注意, 我们预先不知道这个 E_T 的值). 乘上了收敛因子的新 Green 函数当然已经不再是原方程 (22) 的解, 而是存在着一个移位:

$$\frac{\partial \rho(\mathbf{x}, \tau)}{\partial \tau} = \frac{1}{2} \nabla^2 \rho(\mathbf{x}, \tau) - (V - E_T) \rho(\mathbf{x}, \tau). \quad (28)$$

如果我们选取 E_T 使得 Green 函数 (27) 归一化, 那么方程 (28) 就会存在一个不随 τ 演化的解:

$$E_T \rho(\mathbf{x}) = -\frac{1}{2} \nabla^2 \rho(\mathbf{x}) + V(\mathbf{x}) \rho(\mathbf{x}). \quad (29)$$

这时我们就会发现, 式 (29) 正是定态方程 (1). 那么接下来我们就需要求解 (29).

接下来我们考虑 Green 函数:

$$G = \langle \mathbf{x}' | \exp(-\tau(H - E_T)) | \mathbf{x} \rangle \quad (30)$$

$$\approx \frac{\exp(-\tau(V - E_T)) \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2\pi\tau}\right)}{\sqrt{2\pi\tau}} \quad (31)$$

的性质. 我们先把算符 $\exp(-\tau(H - E_T))$ 按它的本征态 $|\phi_n\rangle$ 展开:

$$\begin{aligned} & \exp(-\tau(H - E_T)) \quad (32) \\ &= \sum_n |\phi_n\rangle \exp(-\tau(E_n - E_T)) \langle \phi_n|. \end{aligned}$$

在 τ 很大的情况下基态 $|\phi_0\rangle, E_0$ 在上式以 $\exp[-\tau(E_1 - E_G)]$ 的优势占据主导地位. 所以算符 (32) 起到了一个将任意初态投影到基态的投影算符. 而由于我们只有形如式 (31) 的在 τ 很小的情况下的 G 的近似表达式, 我们必须多次将 Green 函数作用在波函数 $\psi^{(i)}(\mathbf{x})$ 上. 这就是投影 *Monte Carlo*.

在实际的计算模拟中, 一批粒子被模拟在位形空间中进行随机游走, 每个游走过程分为扩散和分支两步. 在扩散步中, 每个粒子按 Green 函数的扩散项进行随机取点 (由于 Gauss 函数的抽样技术非常简单, 这里应该直接使用解析的抽样公式); 而在分支步中, 每个粒子按势能

项 (权函数) 之比的概率

$$\exp\left[\frac{-(V(\mathbf{x}^{(i+1)}) - E_T)}{-(V(\mathbf{x}^{(i)}))}\right]$$

来决定是否真的游走到所取的点 $\mathbf{x}^{(i+1)}$. 这个过程效率不是很高, 因为粒子很容易随机游走到势能权重非常低的地方去, 停留在这类地方会造成计算资源的浪费. 为了缓解这个问题, 我们可以引入“出生-死亡法”: 每当粒子游走到一个点 $\mathbf{x}^{(i+1)}$, 我们计算 $\mathbf{x}^{(i+1)}$ 的权函数:

$$q = \exp[-\tau(V(\mathbf{x}^{(i+1)}) - E_T)].$$

如果 $q < 1$, 这个粒子就有 $1-q$ 的几率“死亡”; 而如果 $q > 1$, 这个粒子就会使 $\text{floor}(q - 1)$ 或 $\text{floor}(q)$ 个粒子在 $\mathbf{x}^{(i+1)}$ 点“出生”, 出生 $\text{floor}(q)$ 个粒子的几率由 $q - \text{floor}(q)$ 给出.

我们注意到, “出生-死亡法”会使得进行随机游走的全部粒子数发生变化, 而 $E_T = E_0$ 时随机游走的粒子数目在统计上是恒定的. 所以我们可以进行了一段时间的模拟之后对 E_T 作如下调整:

$$E_T = E_0 + \alpha \log(\tilde{M}/M),$$

式中 \tilde{M} 是我们期望的粒子数而 M 是真实观察到的粒子数. 在经过一段时间的调整之后, 最终 E_T 的平均值会收敛到系统的基态能量 E_0 , 而随机游走的粒子则会收敛到系统的基态波函数.

投影 QMC 有着比变分 QMC 更好的精度, 而且不依赖于猜测算法的选取. 算法最主要的系统误差来源于被舍弃的 CBH 对易子, 而这也是可以通过缩短 τ 来改善的. 但是投影 QMC 的收敛速度依然严重依赖初始的试探波函数的选取, 这是在编写程序时需要注意的.