

一套重复不出来的力场参数

State: PreAccept

Date of Manuscript: 2026-07-19

Date of PreAccept: 2026-07-31

License: CC BY 4.0李永乐^{1*}¹ 上海大学物理系,上海市宝山区上大路 99 号 E 楼 121, 200444

事情还要从 2009 年左右说起 [1]。这时候台湾中研院的 Carmay Lim (林小乔) 课题组发表了一篇文章, 提到一个能够模拟金属蛋白的力场方案: 不需要对配合物引入额外的化学键, 让电荷随着时间变化, 并引入一个包含电极化率的补偿项 (见附录), 即可模拟金属蛋白中, 金属离子的配位数以及结构随时间的变化。后边还提出了一套迭代公式, 可以通过自洽迭代的办法, 确定这个极化项中的参数。总体思路直接, 非常符合同行们的心理预期: 运用包含极化的电场高阶修正, 来补偿电荷转移带来的能量减小。

1 正文

本文是 2005 年他们两人合作的一篇高水平论文 [2] 的推广。当时我们课题组, 正在基于我们自己的 MFCC-PCC 方法 [3], [4], 模拟金属蛋白。这里边遇到了一些问题, 即这套方案具有强烈的初值依赖性, 无法正确模拟金属的配位数随时间的变化。当时我的导师提出, 可以尝试浮动电荷方案, 即想办法让电荷随时间变化。但如此以来, 每一步 MD 都要做 *ab initio* 量子化学计算, 耗时较多。而且电荷变化之后, 由于金属蛋白中, 金属离子-配体电荷转移一般都是金属向配体转移部分电荷, 金属电量减少, 导致金属离子与配体的相互作用减弱。这个减弱有时候是好事: 可以让金属离子与配体的几何结构动态变化, 更加符合实验中测量到的情况; 但是这个减弱也会导致金属离子的行为变得更像周围的钠离子或者钾离子: 飘向溶液, 离开金属蛋白的配位中心。也就是说静电相互作用减少过多, 需要考虑引入其他能量项, 进行补偿。当时我们在考虑引入某种补偿方案, 又可望比流行的一种可极化力场: AMOEBA 力场 [5] 简单一些。林老师课题组这个方案可以说是想睡觉时有人递枕头。

可惜好景不长, 我们在 2010 年计划测试这套方法, 又跟我一位师兄背靠背各自按照文中方法编写了程序, 发现我俩的程序按照文章的流程计算, 得到的数值均不收敛。我忍不住, 写信给林老师, 她告诉我那个一作俄国人 Sakharov 写完这两篇文章就拿博士学位回国了。她只能说把邮件转给他。看来, 林老师对这个方法的细节也没有

深究。而且考虑到我们的电荷方案跟林老师采用的线性变化电荷也不一样, 干脆不再管。我就以固定 MFCC 电荷的 MD 模拟一段时间, 再根据结构更新一遍电荷的方法, 成功模拟完一个比较有趣的双锌蛋白之后, 拿博士学位走人 [6]。我的师兄也没再继续关注这类体系。当年 2011 年, 加上我们课题组, 有四个课题组发表了金属蛋白模拟的新方法, 其中另外三组分别是我的博后导师 Yingkai Zhang、佛罗里达的 Ken Merz, 以及法国的 N. Gresh。其中后两个课题组长期进行这方面的开发, 虽然 2011 年的时候他们的工作有必须在一开始引入配位键信息、模拟中无法研究配位数变化的问题, 但在直到今天的工作中, 这俩课题组都发展出了新方法克服这一问题。后来我一位师弟采用机器学习的方法, 在 2018 年左右, 又重新做了金属蛋白的模拟, 把我毕业之后遗留的体系做完, 也算是了却一桩多年心愿。因为我毕业之后从事计算化学物理研究, 渐渐脱离了生物体系的模拟。

回到那两篇“系列文章”, 在 2012 年 AMOEBA 力场开发者任鹏宇的综述文章中还获得引用 [7]。但这篇文章后来就渐渐退出历史舞台, 每年引用量仅一两次, 最多有一年达到 4 次。实际上此文引用量最大的时间正好是 2010、2011 年两年, 每年都有十几次引用, 我们相当于正好踩在坑里。

后续直到 2024 年, 林老师课题组又发文章, 继续拓展了这一方法 [8] 还给它取了一个名字 CTPOL。但文章中没有提及自洽迭代, 且引入了新的截断参数。所以说, 这个方法当年的计算细节的正确性, 也就是参数的计算可能

仍然存疑。不过现在至少有代码可查（2024 年文章提供了 GitHub 链接），验证起来更为容易。有兴趣的读者可以自行尝试复现这篇文章中的计算。这也是目前的业界普遍的规范：论文应提供完整的代码、参数、数据、测试算例和收敛标准，否则公式即使再完整，也未必足以保证可复现、可迁移。希望以后不再有博士生们，再遭受我们当年的痛苦。

2 附录

浮动的电荷改变量： $\Delta q(t) = ar(t) + b$ ，其中 $r(t)$ 是随时间变化的金属-配体距离， a 和 b 是拟合常数（按照今天的流行说法，就是待训练参数）。然后在分子动力学（molecular dynamics, MD）模拟时，让电荷随时间变化： $q(t) = q + \Delta q(t)$ ；再补上偶极矩-外场相互作用极化（polarization）项：

$$E^{\text{pol}} = -\frac{1}{2} \sum_i \vec{\mu}_i \cdot \vec{E}_i^0 = -\frac{1}{2} \sum_i \alpha_i \vec{E}_i \cdot \vec{E}_i^0$$

这里 μ_i 是第 i 个原子的偶极矩，待定。上标 0 代表电荷部分，而不带 0 的电场强度包括诱导偶极的贡献：

$$\begin{aligned} \vec{E}_i &= \vec{E}_i^0 + \sum_j \hat{T}_{ij} \vec{\mu}_j \\ &= \sum_j \frac{q_j \vec{r}_{ij}}{r_{ij}^3} + \sum_j \frac{1}{r_{ij}^3} \left(\frac{3\vec{r}_{ij} \vec{r}_{ij}}{r_{ij}^2} - \hat{1} \right) \vec{\mu}_j \end{aligned}$$

第一行公式中，加帽子的 $\hat{T}$ 表示矩阵。第二行公式中，右边的两个矢量写在一起表示并矢。加帽子的 $\hat{1}$ 表示这里是单位矩阵。这样，有

$$\begin{aligned} \vec{\mu}_i &= \alpha_i \vec{E}_i(\{\vec{\mu}_i\}) \\ &= \alpha_i \left\{ \sum_j \frac{q_j \vec{r}_{ij}}{r_{ij}^3} + \sum_j \frac{1}{r_{ij}^3} \left(\frac{3\vec{r}_{ij} \vec{r}_{ij}}{r_{ij}^2} - \hat{1} \right) \vec{\mu}_j \right\} \end{aligned}$$

也就是指定一套初始的偶极矩，可望通过自洽迭代得到收敛的偶极矩数值。但我跟我师兄的尝试计算，均显示该偶极矩不收敛。

参考文献

- [1] D. V. Sakharov and C. Lim, *Journal of Computational Chemistry*, vol. 30, p. 191, 2009.
- [2] D. V. Sakharov and C. Lim, *Journal of the American Chemical Society*, vol. 127, p. 4921, 2005.
- [3] C. Ji, Y. Mei, and J. Zhang, *Biophysical Journal*, vol. 95, p. 1080, 2008.
- [4] C. Ji and Y. Mei, *Accounts of Chemical Research*, vol. 47, p. 2795, 2014.
- [5] R. Ren and J. Ponder, *Journal of Physical Chemistry B*, vol. 107, p. 5933, 2003.
- [6] Y. Li, Y. Mei, D. Zhang, D. Xie, and J. Zhang, *Journal of Physical Chemistry B*, vol. 115, p. 10154, 2011.
- [7] W. Yang, P. Ren, and others, *Journal of Chemical Theory and Computation*, vol. 8, p. 1314, 2012.
- [8] X. Hu, C. Lim, and others, *Journal of Chemical Theory and Computation*, vol. 20, p. 1448, 2024.

编辑评论

这篇文章反映了早期计算科学论文中一个并不少见的问题：文章给出了公式和结果，却没有充分说明参数如何选择、迭代如何实现、数值为何能够收敛。许多关键细节隐藏在作者自己的程序和经验中，论文表面上看似完整，真正照着做时却未必能够重复。

仅凭一次或两次复现失败，当然不能直接断定原作者存在造假；究竟是方法本身有问题、论文遗漏了关键步骤，还是实现条件不同，目前仍然存疑。但可以确定的是，这种写得不清楚、解释不充分、又缺少代码和测试数据的文章，很难接受真正的检验。

科学研究不仅要给出一个看起来正确的结果，也应当让其他人明白这个结果是怎样得到的。否则，一套方法即使发表在高水平期刊上，也可能只是一个无法核查的“黑箱”。这样的论文不仅会浪费后来者的时间，也会阻碍知识的积累和科学的发展。