Benjamin Hill

Brother Martin

MATH 412

April 8, 2024

Molecular Conformation

The protein folding problem has become a hotbed of multidisciplinary optimization research. Simulated annealing and powerful quasi-Newton Methods are often used to predict conformation of complicated molecules, with ever more realistic modeling of the intermolecular forces. The Protein Data Bank is a useful worldwide archive of structural data on biological macromolecules. Extensive lists of experimentally measured atom positions are available there for use. Here we will explore some ways these values can be calculated, comparing the methods involved. Note that I collaborated with Danen and Joe to develop the code for this exploration.

Part 1

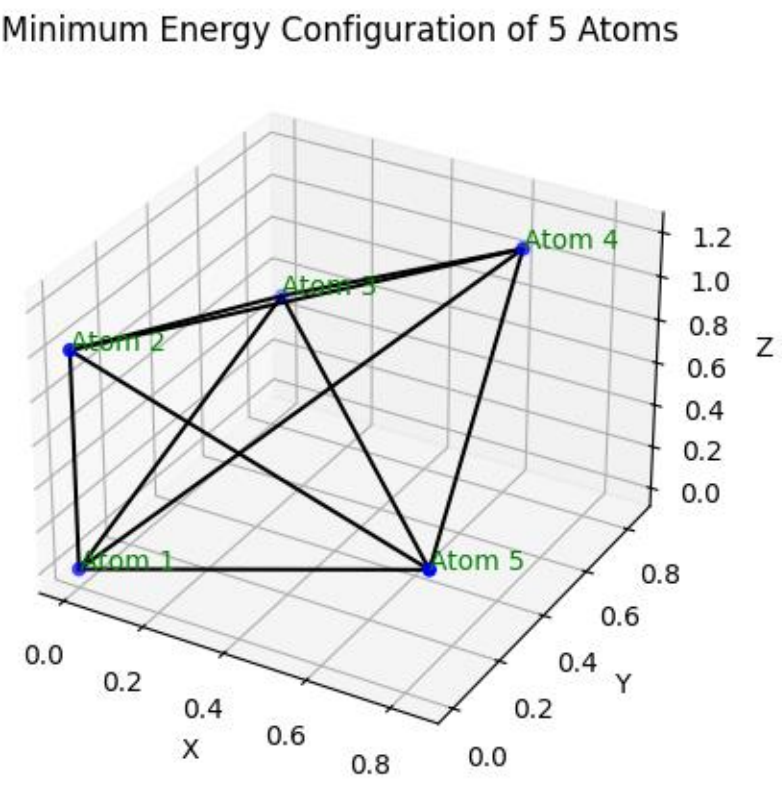
We begin with the model for potential energy between atoms. For our purposes here, we will use the following formula for potential energy U:

$$U = \sum_{i<j} \left(\frac{1}{{r_{ij}}^{12}} - \frac{2}{{r_{ij}}^{6}} \right)$$

where r_{ij} is the Euclidean distance between atoms i and j. Given a set of atoms at certain coordinates, using this formula between each pair of atoms gives the potential energy of the position. Naturally, the atoms will be drawn toward an equilibrium where they are neither attracted nor repelled by the other atoms, such that all of the atoms are stable in the absence of other disturbances. Thus, we can use minimizing processes to find this equilibrium energy and the position of the atoms in that state. Starting with a set of five atoms at random locations, we use the Nelder-Mead optimizing algorithm to find the minimum energy. This process yields a potential energy of -8.67616973, which the algorithm took 664 steps to reach.

Part 2

The resulting coordinates of the atoms from Part 1 are depicted below.



Part 3

Now we will develop a gradient-based method for optimizing the potential energy. We do this by approximating the gradient at the coordinates of each atom and then adjusting the atoms slightly in the direction of lowest gradient. Then we measure the again and repeat the process until there is very little change between one iteration and the next. Using this logic, we implement the conjugate gradient method and analyze the position again. This time, we find a minimum potential energy of -9.10385242 between the atoms.

Part 4

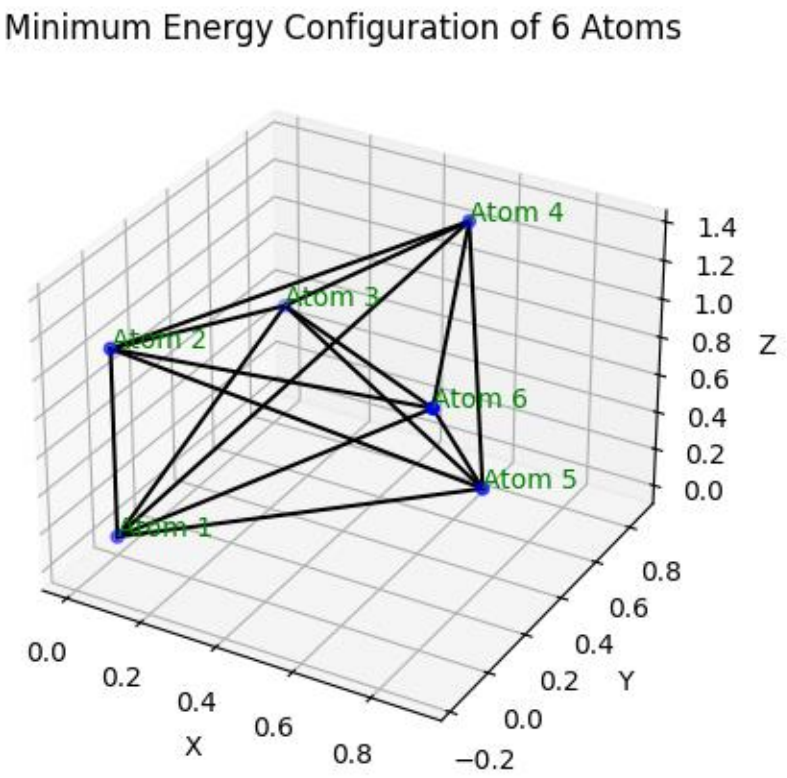
Using a different non-gradient method to check our numbers would be helpful, so we implement the Powell optimization method to find the global minimum of the potential energy between the five atoms. This yields an energy of -9.10315754, which is very close to the value given by the conjugate gradient method.

Part 5

Now we will try this again with six atoms instead of five. Going through the same analysis as before, Nelder-Mead gives an energy of -11.24408635 in 1745 steps, while the conjugate gradient method gives an energy of -12.30292753 in only 101 iterations. As before, the global minimum according to Powell's method is -12.29912296, much closer to the result of the conjugate gradient method than to Nelder-Mead, and the conjugate gradient method gets to that higher accuracy more than ten times as quickly. From these results, it is clear that the conjugate gradient method is far superior to Nelder-Mead, at least for this particular problem.

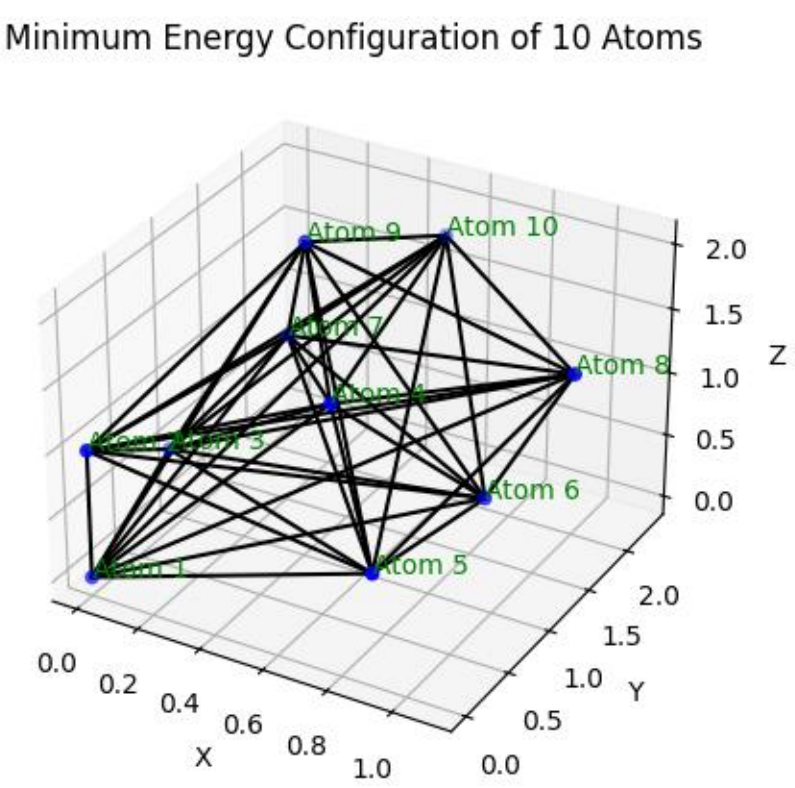
Part 6

The resulting coordinates of the atoms from Part 5 are depicted below.



Part 7

Now we will determine the minimum energy and corresponding atom positions for a larger number of atoms. For now we will only explore a set of ten atoms. The conjugate gradient method yields a minimum energy of -25.33781177, which is a trustworthy value, according to our previous experiences with this method. The coordinates of this minimum-energy state are shown below.


Minimum Energy Configuration of 10 Atoms
Atom 9
Atom 10
Atom 8
Atom 6
Atom 5
0.0
0.2
0.4
0.6
0.8
1.0
X
0.0
0.5
1.0
1.5
2.0
Y
0.0
0.5
1.0
1.5
2.0
Z