Cryo-EM Heterogenous conformations

Claire Donnat, Yuehaw Khoo, Sounak Paul

April 2022

1 Algorithm

The following is the setting of the problem. Let,

- F be the function putting gaussian blobs on 2D projections of rotated coordinates.
- X_0 be the given initial template of atomic coordinates.
- $I_1, ..., I_n$ are the n given cryo-em images.
- $R_1, ..., R_n$ are the n given rotations.
- $X_1, ..., X_n$ are the n heterogeneous conformations at n different timestamps that we need to find.

Now, let $\mu(X)$ and $\nabla\mu(X)$ be the potential energy and the forces of the state X. Then, we know that

$$p(X) \propto e^{-\frac{\mu(X)}{k_B T}}.$$

Also, let

$$X_i \;=\; X_0 + MLP_\theta(X_i - X_0),$$

where we wish to learn θ. Assuming I_i to be a gaussian r.v. centered around X_i, we get the following negative log-likelihood loss function,

$$\begin{aligned} L \;&=\; -\log(p(X_i)) \\ &\propto\; -\log(p(I_i|X_i)) - \log(p(X_i)) \\ &=\; \frac{||I_i - F(X_i)||^2}{2\sigma^2} + \frac{\mu(X_i)}{k_B T} \\ \therefore\; \frac{\partial L}{\partial \theta} \;&=\; \frac{F(X_i) - I_i}{\sigma^2}\frac{\partial F(X_i)}{\partial X_i}\frac{\partial X_i}{\partial \theta} \;+\; \frac{\nabla\mu(X)^T}{k_B T}\frac{\partial X_i}{\partial \theta} \end{aligned}$$

Hence, the updates become

$$\theta_{t+1} \;=\; \theta_t \;-\; \eta\frac{\partial L}{\partial \theta}.$$