Basic idea to reformulate Recitfied Linear Unit (ReLu)

$$ReLu(x) = ln(1 + exp(x))$$

In an MLP, from one layer to the next, each node activation $\mathbf{z} = \mathbf{Wh} + \mathbf{b}$

WLOG, let the bias $\mathbf{b} = 0$. Also, assume ReLu is applied elementwise to the vector $\mathbf{z}$

Then, for each activated node $\mathbf{n}_i$

$$n_i = ReLu(\mathbf{z}_j) = ln(1 + exp((\mathbf{Wh})_j)$$

$$= ln(1 + exp((\sum_i \mathbf{W_{ij}h_i}))$$

$$= ln(1 + \prod_i exp((\mathbf{W_{ij}h_i}))$$

Which seems intractable

But

What if want consider $Relu(\mathbf{z}_j)$ in Expectation

$$\mathbb{E}[Relu(\mathbf{z}_j)]$$

we assume that each activated node exists in a superposition of 2 states [0,1]

what I want to do is somehow get the superposition inside the log, so that

$$\mathbb{E}[Relu(\mathbf{z}_j)] \approx ln(\mathbb{E}[1 + exp((\mathbf{Wh})_j])$$

this is kindof like the annealed approximation in spin glass theory–which is not a great approximation, but perhaps, can give some initial insight

from here, lets introduce an internal state vector $\mathbf{c} = [0, 1]$

$$ln(\mathbb{E}[1 + exp((\mathbf{Wh})_j]) = ln(\mathbb{E}[exp((\mathbf{c^T Wh})_j])$$

and we can get something like

$$\mathbb{E}[Relu(\mathbf{z}_j)] \approx ln(\mathbb{E}[exp((\mathbf{c^T Wh})_j])$$

We now pull the expecation back out...

$$\approx \mathbb{E}[ln(exp((\mathbf{c^T Wh})_j)]$$

$$\approx \mathbb{E}[(\mathbf{c^T Wh})_j]$$

So obviously this is very sloppy, but it is an attempt to get at defining some kind of (linearized) free energy for each layer