1 Method

I trained a Restricted Boltzmann Machine (RBM) on the XOR dataset for 3-bit inputs. Of the 8 combinations, 4 patterns output +1 with probability 0.25, while the others have 0 probability. The goal is to train a network that generates 3-bit patterns with $P_{\text{Boltzmann}}$ replicating P_{data}. I explored various M values, expecting 8 neurons to be sufficient, with 4 neurons yielding good Kullback-Leibler Divergence results. The network employs the Contrastive Divergence (CD-k) algorithm, with 3 visible neurons, M hidden neuron and parameters $v_{\max} = 10000$, $p_0 = 20$, and $\eta = 0.005$. I iterate the dynamics of the trained Boltzmann machine 100,000 times, counting occurrences of patterns with $P_{\text{data}} \neq 0$, representing model probabilities. I train 20 different Boltzmann machines for each M and compute divergences between data probabilities and observed probabilities (Eq. 1).

$$D_{KL} = \sum_{\mu=1}^{p} P_{\text{data}}(x(\mu)) \log \left[\frac{P_{\text{data}}(x(\mu))}{P_B(s = x(\mu))} \right] \tag{1}$$

2 Results

I compute the upper bound for each divergence score (Eq. 2) and plot the scores of the 20 simulations for each M alongside KL divergence upper bounds (Fig. 1).

$$D_{KL} \leq \begin{cases} \log(2) \left(N - \lfloor \log_2(M+1) \rfloor - \frac{M+1}{2^{\lfloor \log_2(M+1) \rfloor}} \right) & \text{if } M < 2^{N-1} - 1 \\ 0 & \text{if } M \geq 2^{N-1} - 1 \end{cases} \tag{2}$$

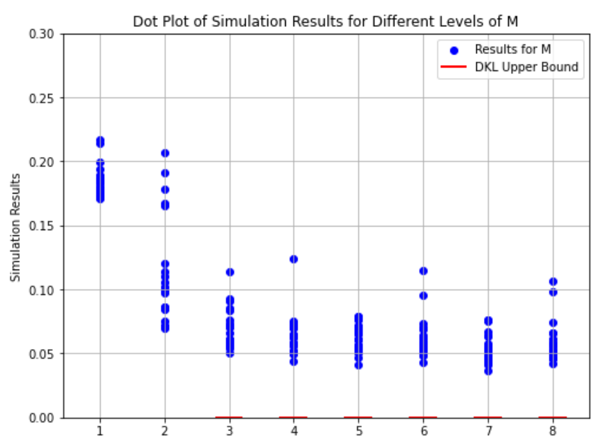

Figure 1: Divergence scores and KL divergence upper bounds for varying M.

The initial KL bounds of 0.69 and 0.34 are not visible as the plot cuts at 0.30 in order to improve visibility. For $M = 1, 2$, simulated divergences are below upper bounds; for $M > 2$, they remain above. In the case of $M = 1, 2$ the divergences are more sensitive to parameter choices as in some instances I observed divergences approach 0.50 and 0.30, indicating a larger margin for error. However, for $M > 2$, results are consistently low, suggesting it is easier to fit the problem with more hidden neurons. The mean of the 20 simulations decreases and becomes less dispersed as M increases, yet I do not achieve perfect zero divergence. Simulated probabilities approach 0.25 but do not always sum to 1, indicating the possibility of generating unseen patterns with low a probability.