Calo4pQVAE: Progress and updates

Nov 28 2024

RBM to Diffusion Model equivalence

Let's consider a two-partite RBM:

$$Z = \sum_{v,h} e^{-\beta E(v,h)}$$

$$E(v,h) = -\langle v | a_0 \rangle - \langle b_0 | h \rangle - \langle v | W | h \rangle$$

The vectors v and h are binary vectors, i.e., $v_i \in \{0,1\}$ and $h_i \in \{0,1\}$

We'll assume as usual that the visible layer has N nodes and the binary layer has M nodes with N>M

Problem: If N+M >~ 50, computing the partition function Z becomes intractable.

RBM to Diffusion Model equivalence

Let's apply Singular Value Decomposition (SVD)

$$Z = \sum_{v,h} e^{-\beta E(v,h)}$$

$$E(v,h) = -\langle v \,|\, a_0 \rangle - \langle b_0 \,|\, h \rangle - \langle v \,|\, W \,|\, h \rangle$$

$W = U\Sigma V^t$ (SVD)

$$UU^t = U^tU = I$$

$$VV^t = V^tV = I$$

Σ is a rectangular matrix NxM and contains the singular values, λ's, in the diagonal

RBM to Diffusion Model equivalence

Let's apply Singular Value Decomposition (SVD)

$$Z = \sum_{v,h} e^{-\beta E(v,h)}$$

$$E(v,h) = -\langle v|a_0\rangle - \langle b_0|h\rangle - \langle v|W|h\rangle$$

$W = U\Sigma V^t$ (SVD)

$$UU^t = U^tU = I$$

$$VV^t = V^tV = I$$

Σ is a rectangular matrix NxM and contains the singular values, λ's, in the diagonal

We define the new variables in reciprocal space

$$|x\rangle = U|v\rangle, \qquad |y\rangle = V|h\rangle$$

$$E(v,h) = -\langle v|a_0\rangle - \langle b_0|h\rangle - \langle v|W|h\rangle$$

$\uparrow$ $I = U^tU$ $\quad$ $\uparrow$ $I = V^tV$ $\quad$ $\uparrow$ $U\Sigma V^t$

$$E(x,y) = -\langle x|a\rangle - \langle b|y\rangle - \langle x|\Sigma|y\rangle$$

$$Z = \sum_{x,y} e^{-\beta E(x,y)}$$

RBM to Diffusion Model equivalence

We don't get much accomplished by expressing the partition function in terms of x's and y's :(

$$E(x, y) = -\langle x | a \rangle - \langle b | y \rangle - \langle x | \Sigma | y \rangle$$

$$Z = \sum_{x,y} e^{-\beta E(x,y)}$$

X's and y's are discrete, since each vector is a linear and unitary transformation of v and h

$$|x\rangle = U|v\rangle, \qquad |y\rangle = V|h\rangle$$

RBM to Diffusion Model equivalence

Let's relax the constraints on x's and y's. Let us assume the x's and y's are continuous variables

$$Z = \sum_{v,h} e^{-\beta E(v,h)} = \sum_{v,h} \prod_i \prod_j \iint dx_i dy_j e^{-\beta E(x,y)} \delta\left(x_i - \sum_k U^t_{ik} v_k\right) \delta\left(y_j - \sum_l V^t_{jl} h_l\right)$$

Constraints on x's and y's. This term will *pick up* the values of x's and y's that match the SVD transformation

$$Z = 2^{N+M} \iint dx_i dy_j e^{-\beta E(x,y)} \rho(\{x\}, \{y\})$$

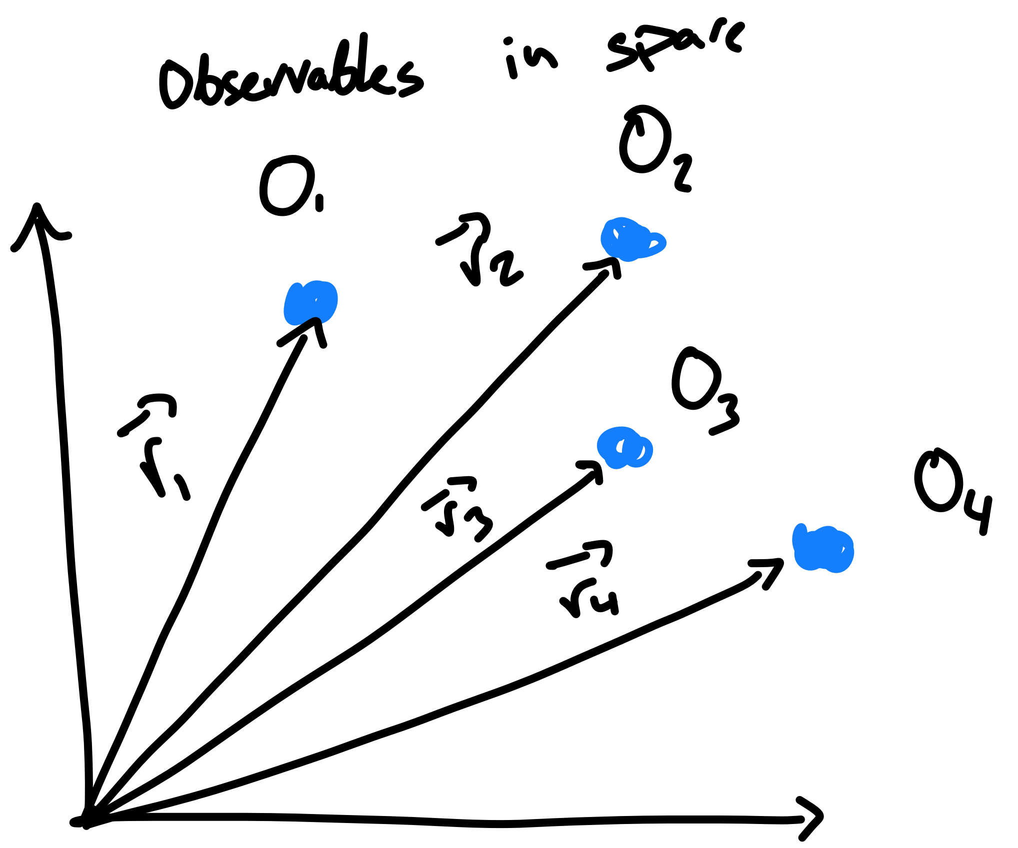
$W = U \Sigma V^t$ (SVD)

$$|x\rangle = U|v\rangle, \qquad |y\rangle = V|h\rangle$$

RBM to Diffusion Model equivalence

$$\rho(\boldsymbol{x},\boldsymbol{y}) \equiv \langle \prod_{i=1}^{N}\prod_{j=1}^{M} \delta(\sum_k U^t_{ik} v_k - x_i)\delta(\sum_l V^t_{jl} h_l - y_j)\rangle_{(\boldsymbol{v},\boldsymbol{h})}$$

$$= \frac{1}{2^{N+M}} \sum_{\{\boldsymbol{v},\boldsymbol{h}\}} \prod_{i=1}^{N}\prod_{j=1}^{M} \delta(\sum_k U^t_{ik} v_k - x_i)\delta(\sum_l V^t_{jl} h_l - y_j)$$



mean value: $\langle O \rangle = \frac{1}{4}(O_1 + O_2 + O_3 + O_4)$

We can also define a density

$$\rho(\vec{r}) = \frac{1}{4}\sum_{i=1}^{4} \delta(\vec{r} - \vec{r}_i)$$

such that

$$\langle O \rangle = \int d^3\vec{r}\, \rho(\vec{r})\, O(\vec{r})$$

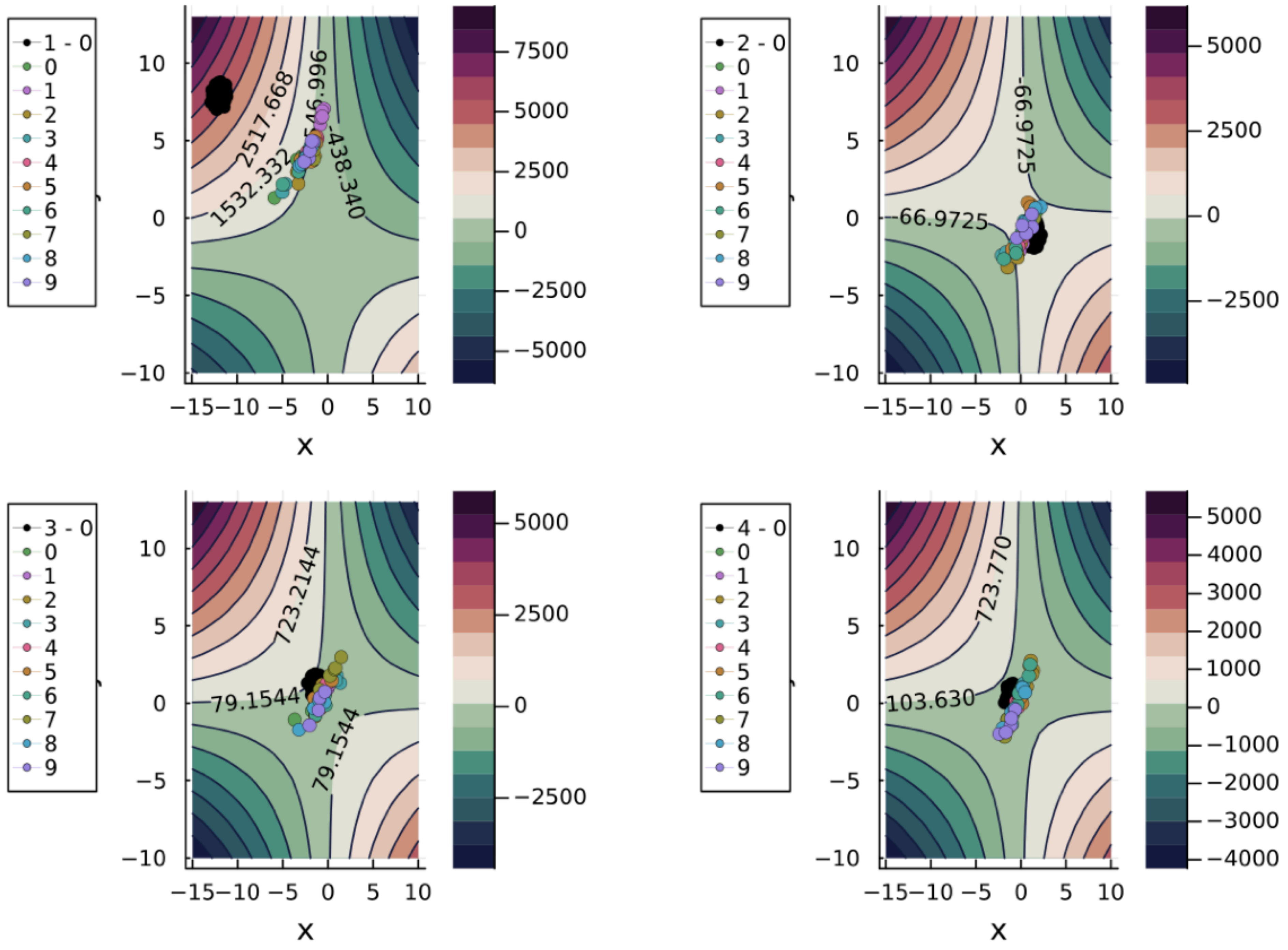
$$= \frac{1}{4}\sum_{i=1}^{4} O(\vec{r}_i) = \frac{1}{4}\sum_{i=1}^{4} O_i$$

RBM to Diffusion Model equivalence

Short pause here!

$$E(x,y) = -\langle x|a\rangle - \langle b|y\rangle - \langle x|\Sigma|y\rangle$$

$$E_i = \begin{cases} -a_{0i}x_i - b_{0i}y_i - \lambda_i x_i y_i, & 1 \leq i \leq min(N,M) \\ -a_{0i}x_i, & min(N,M) \leq i \leq max(N,M) \end{cases}$$



The energy landscape in reciprocal space are hyperbolas

We can align the principal axis with the reference system by shifting to the saddle points and rotating pi/4 radians

RBM to Diffusion Model equivalence

$$\rho(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{x}) \equiv \langle \prod_{i=1}^{M} \prod_{j=M+1}^{N} \delta(u_i - \frac{1}{\sqrt{2}}(\sum_k U_{ik}^t v_k - \sum_k V_{ik}^t h_k - x_{0i} + y_{0i}))$$
$$\delta(w_i - \frac{1}{\sqrt{2}}(\sum_k U_{ik}^t v_k + \sum_k V_{ik}^t h_k - x_{0i} - y_{0i}))\delta(\sum_k U_{jk}^t v_k - x_j)\rangle_{(\boldsymbol{v},\boldsymbol{h})}$$
$$= \frac{1}{2^{N+M}} \sum_{\{\boldsymbol{v},\boldsymbol{h}\}} \prod_{i=1}^{M} \prod_{j=M+1}^{N} \delta(u_i - \frac{1}{\sqrt{2}}(\sum_k U_{ik}^t v_k - \sum_k V_{ik}^t h_k - x_{0i} + y_{0i}))$$
$$\delta(w_i - \frac{1}{\sqrt{2}}(\sum_k U_{ik}^t v_k + \sum_k V_{ik}^t h_k - x_{0i} - y_{0i}))\delta(\sum_k U_{jk}^t v_k - x_j)$$

This is a really scary term! We're just going to box it in as ρ

RBM to Diffusion Model equivalence

Finally! A partition function of an energy based models with continuous variables!

$$z = \left[u_{1:M}, w_{1:M}, x_{M+1:N}\right]^t$$

$$Z = \int d\boldsymbol{z} e^{-\beta U_{eff}(\boldsymbol{z})}$$

$$U_{eff}(\boldsymbol{z}) = E(\boldsymbol{z}) - \frac{S_c}{\beta} + \frac{V_{const}(\boldsymbol{z})}{\beta}$$

$$V_{const}(\boldsymbol{z}) = -\ln \rho(\boldsymbol{z})$$

$$S_c = 2^{N+M}$$

$$P(\boldsymbol{z}) = \frac{e^{-\beta U_{eff}(\boldsymbol{z})}}{Z}$$

This must be the stationary solution of a Fokker-Planck Equation!

RBM to Diffusion Model equivalence

Finally! A partition function of an energy based models with continuous variables!

$$P(\boldsymbol{z}) = \frac{e^{-\beta U_{eff}(\boldsymbol{z})}}{Z}$$

This must be the stationary solution of a Fokker-Planck Equation!

$$\frac{\partial P(\boldsymbol{z},t)}{\partial t} = D \sum_{i=1}^{N+M} \left[\frac{\partial^2 P(\boldsymbol{z},t)}{\partial z_i^2} + \frac{1}{D} \frac{\partial}{\partial z_i} \left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i} P(\boldsymbol{z},t) \right) \right]$$

F-P Eq in the over damped regime with friction coefficient =1

$$D = 1/\beta$$

Fluct-Diss Theo

RBM to Diffusion Model equivalence

Finally! A partition function of an energy based models with continuous variables!

$$P(\boldsymbol{z}) = \frac{e^{-\beta U_{eff}(\boldsymbol{z})}}{Z}$$

This must be the stationary solution of a Fokker-Planck Equation!

$$\frac{\partial P(\boldsymbol{z},t)}{\partial t} = D \sum_{i=1}^{N+M} \left[\frac{\partial^2 P(\boldsymbol{z},t)}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i} P(\boldsymbol{z},t) \right) \right]$$

F-P Eq in the over damped regime with friction coefficient =1

$$D = 1/\beta$$

Fluct-Diss Theo

Let's try to solve it!

RBM to Diffusion Model equivalence

$$\frac{\partial P(\boldsymbol{z},t)}{\partial t} = D\sum_{i=1}^{N+M}\left[\frac{\partial^2 P(\boldsymbol{z},t)}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}P(\boldsymbol{z},t)\right)\right]$$

- ✦By separation of variables $P(z,t) = f(t)\phi(z)$

$$\frac{\dot{f}(t)}{f(t)} = \frac{D}{\phi}\sum_{i=1}^{N+M}\left[\frac{\partial^2 \phi(\boldsymbol{z})}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}\phi(\boldsymbol{z})\right)\right] = -\Gamma$$

$$f(t) = Ae^{-\Gamma t}$$

RBM to Diffusion Model equivalence

$$\frac{\partial P(\boldsymbol{z},t)}{\partial t} = D\sum_{i=1}^{N+M}\left[\frac{\partial^2 P(\boldsymbol{z},t)}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}P(\boldsymbol{z},t)\right)\right]$$

✦By separation of variables $P(z,t) = f(t)\phi(z)$

$$\frac{\dot{f}(t)}{f(t)} = \frac{D}{\phi}\sum_{i=1}^{N+M}\left[\frac{\partial^2 \phi(\boldsymbol{z})}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}\phi(\boldsymbol{z})\right)\right] = -\Gamma$$

$$f(t) = Ae^{-\Gamma t}$$

✦Let us use the ansatz $\phi(z) = e^{-U_{eff}(z)/2D}\psi(z)$

RBM to Diffusion Model equivalence

$$\frac{\partial P(\boldsymbol{z},t)}{\partial t} = D\sum_{i=1}^{N+M}\left[\frac{\partial^2 P(\boldsymbol{z},t)}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}P(\boldsymbol{z},t)\right)\right]$$

✦By separation of variables $P(z,t) = f(t)\phi(z)$

$$\frac{\dot{f}(t)}{f(t)} = \frac{D}{\phi}\sum_{i=1}^{N+M}\left[\frac{\partial^2\phi(\boldsymbol{z})}{\partial z_i^2} + \frac{1}{D}\frac{\partial}{\partial z_i}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}\phi(\boldsymbol{z})\right)\right] = -\Gamma$$

$$f(t) = Ae^{-\Gamma t}$$

✦Let us use the ansatz $\phi(z) = e^{-U_{eff}(z)/2D}\psi(z)$

$$H\psi(z) = E\psi(z)$$

$$H = \sum_{i=1}^{N+M}\frac{p_i^2}{2m} + V_Q^{(i)}$$

$$V_Q^{(i)}(\boldsymbol{z}) = \frac{1}{8D^2}\left(\frac{\partial U_{eff}(\boldsymbol{z})}{\partial z_i}\right)^2 - \frac{1}{4D}\frac{\partial^2 U_{eff}(\boldsymbol{z})}{\partial z_i^2}$$

RBM to Diffusion Model equivalence

✦We can write down the formal solution:

$$P(\boldsymbol{z},t) = \underbrace{c_0 e^{-\frac{1}{2D}U_{eff}(\boldsymbol{z})}\psi_0(\boldsymbol{z})}_{\text{Stationary solution}} + \underbrace{\sum_{\Gamma_n>0} c_{\Gamma_n} e^{-\frac{1}{2D}U_{eff}(\boldsymbol{z})}\psi_{\Gamma_n}(\boldsymbol{z})e^{-\Gamma_n t}}_{\text{Transient states}}$$

Notice that we haven't done any approximation, except assume the z variables to be continuous subject to a constraint potential embedded in U_{eff}

RBM to Diffusion Model equivalence

We can expand $\rho(z)$ around the max-log-likelihood. This is equivalent to a Taylor expansion of V_{const} around the minimum

$$\omega_i = \begin{cases} \frac{k_{ii}-\lambda\beta}{2} & \text{for } i = 1, ..., M \\ \frac{k_{ii}+\lambda\beta}{2} & \text{for } i = M+1, ..., 2M \\ \frac{k_{ii}}{2} & \text{for } i = 2M+1, ..., N \end{cases}$$

$$V_Q^{(i)} = \frac{1}{2}\omega_i^2\zeta_i^2 + \frac{1}{2}\omega_i\zeta_i\sum_{j\neq i}\zeta_j k_{ij} + \frac{1}{2}\omega_i\zeta_i\sum_{j\neq i}(z_{j0}-\mu_{z_j})k_{ij} + \frac{1}{2}\left(\frac{1}{2}\sum_{j\neq i}(\zeta_j + z_{j0} - \mu_{z_j})k_{ij}\right)^2 - \frac{1}{2}\omega_i$$

$$z_{i0} = \begin{cases} \frac{\mu_{z_i}k_{ii}}{2\omega_i} & \text{for } i = 1, ..., 2M \\ \frac{\beta a_i}{2\omega_i} + \mu_{z_i} & \text{for } i = 2M+1, ..., N \end{cases}$$

$$\zeta_i = z_i - z_{i0}$$

The RBM maps to a F-P Eq. Where the transient states satisfy the Schroedinger Eq. Of a set of couples harmonic oscillators!

RBM to Diffusion Model equivalence

We can expand $\rho(z)$ around the max-log-likelihood. This is equivalent to a Taylor expansion of V_{const} around the minimum

$$\omega_i = \begin{cases} \frac{k_{ii}-\lambda\beta}{2} & \text{for } i = 1, \ldots, M \\ \frac{k_{ii}+\lambda\beta}{2} & \text{for } i = M+1, \ldots, 2M \\ \frac{k_{ii}}{2} & \text{for } i = 2M+1, \ldots, N \end{cases}$$

$$V_Q^{(i)} = \frac{1}{2}\omega_i^2\zeta_i^2 + \frac{1}{2}\omega_i\zeta_i\sum_{j\neq i}\zeta_j k_{ij} + \frac{1}{2}\omega_i\zeta_i\sum_{j\neq i}(z_{j0} - \mu_{z_j})k_{ij}$$
$$+\frac{1}{2}\left(\frac{1}{2}\sum_{j\neq i}(\zeta_j + z_{j0} - \mu_{z_j})k_{ij}\right)^2 - \frac{1}{2}\omega_i$$

$$z_{i0} = \begin{cases} \frac{\mu_{z_i}k_{ii}}{2\omega_i} & \text{for } i = 1, \ldots, 2M \\ \frac{\beta a_i}{2\omega_i} + \mu_{z_i} & \text{for } i = 2M+1, \ldots, N \end{cases}$$

$$\zeta_i = z_i - z_{i0}$$

$$E(\boldsymbol{n}) = \sum_{i=1}^{N+M}\omega_i\left(n_i + \frac{1}{2}\right) - \frac{\omega_i}{2} = \sum_{i=1}^{N+M}\omega_i n_i$$

ToDo

- ✓ Upload camera-ready paper and poster to Neurips
- ✓ Looked into estimating partition function using Tensor Cross Interpolation. https://arxiv.org/abs/2407.02454 but it did not work.
- ✓ How does dwave perform when estimating observables (e.g., energy) when the weights and biases are sampled from a Uniform distribution.
- ✓ Once we get the subscription up and going => Generate samples from Dwave for paper
- ✓ There might be a way to validate this RBM to Diffusion Model via the relaxation times.