

# GAUSSIAN-BINARY RESTRICTED BOLTZMANN MACHINES

Notation:

- visible nodes  $\vec{x} \in \mathbb{R}^M$  - gaussian units (position coordinates)
- hidden nodes  $\vec{h} \in \mathbb{R}^N$  - binary units
- visible biases  $\vec{a} \in \mathbb{R}^M$
- hidden biases  $\vec{b} \in \mathbb{R}^N$
- interaction weights  $\hat{W} \in \mathbb{R}^{M \times N}$
- $j$ th column vector of weight matrix  $\vec{W}_j \in \mathbb{R}^M$
- $i$ th row vector of weight matrix  $\vec{W}_i^T \in \mathbb{R}^N$
- variational parameters  $\vec{\theta} = (a_0, \dots, a_{M-1}, b_0, \dots, b_{N-1}, W_{0,0}, \dots, W_{M-1,N-1})$

"Energy" of a configuration of nodes:

$$\begin{aligned} E_{\vec{\theta}}(\vec{x}, \vec{h}) &= \frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 - \sum_{j=0}^{N-1} b_j h_j - \frac{1}{\sigma^2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x_i W_{ij} h_j \\ &= \frac{1}{2\sigma^2} \|\vec{x} - \vec{a}\|^2 - \vec{b}^T \vec{h} - \frac{1}{\sigma^2} \vec{x}^T \hat{W} \vec{h} \end{aligned}$$

Use the marginal probability to represent the wavefunction as a function of the RBM inputs:

$$\begin{aligned} \Psi_{\vec{\theta}}(\vec{x}) &= \frac{1}{Z} \sum_{\vec{h}} e^{-E_{\vec{\theta}}(\vec{x}, \vec{h})} \\ &= \frac{1}{Z} \sum_{\{h_j\}} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 + \sum_{j=0}^{N-1} b_j h_j + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x_i W_{ij} h_j \right) \\ &= \frac{1}{Z} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 \right) \sum_{\{h_j\}} \exp \left( \sum_{j=0}^{N-1} b_j h_j + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x_i W_{ij} h_j \right) \\ &= \frac{1}{Z} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 \right) \prod_{j=0}^{N-1} \sum_{h_j=0}^1 \exp \left( b_j h_j + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} x_i W_{ij} h_j \right) \\ &= \frac{1}{Z} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 \right) \prod_{j=0}^{N-1} \left( 1 + \exp \left( b_j + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} x_i W_{ij} \right) \right) \\ &= \frac{1}{Z} \exp \left( -\frac{1}{2\sigma^2} \|\vec{x} - \vec{a}\|^2 \right) \prod_{j=0}^{N-1} \left( 1 + \exp \left( b_j + \frac{1}{\sigma^2} \vec{x}^T \vec{W}_j \right) \right) \end{aligned}$$

Hamiltonian for  $P$   $D$ -dimensional particles in a harmonic oscillator:

$$\hat{H} = \frac{1}{2} \sum_{p=0}^{P-1} \left( -\nabla_p^2 + \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \right) + \sum_{p < q} V_{int}(r_{pq}), \quad r_{pq} = \sqrt{\sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})^2}$$

We need to vectorize the coordinates of all the particles to input into the RBM:

$$\vec{x} = (r_{0,0}, \dots, r_{0,D-1}, \dots, r_{P-1,0}, \dots, r_{P-1,D-1}) \in \mathbb{R}^M, \quad M = PD$$

The visible biases and the rows of the weight matrix are organized in the same way. The mapping between coordinates and the visible nodes are:

$$\begin{aligned} r_{p,d} &= x_{Dp+d} \\ x_i &= r_{\text{floor}(i/D), i \bmod D} \end{aligned}$$

Now we can write our representation of the wavefunction as a function of all coordinates:

$$\begin{aligned} \Psi_{\vec{\theta}}(\vec{R}) &= \frac{1}{Z} \exp \left( -\frac{1}{2\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} (r_{p,d} - a_{Dp+d})^2 \right) \prod_{j=0}^{N-1} \left( 1 + \exp \left( b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} \right) \right) \\ &= \frac{1}{Z} \exp(A(\vec{R})) \prod_{j=0}^{N-1} \exp \left( \ln \left( 1 + \exp \left( b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} \right) \right) \right) \\ &= \frac{1}{Z} \exp(A(\vec{R})) \exp \left( \sum_{j=0}^{N-1} \ln \left( 1 + \exp \left( b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} \right) \right) \right) \\ &= \frac{1}{Z} \exp \left( A(\vec{R}) + B(\vec{R}) \right) \end{aligned}$$

$$A(\vec{R}) \equiv -\frac{1}{2\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} (r_{p,d} - a_{Dp+d})^2 = -\frac{1}{2\sigma^2} \sum_{i=0}^{M-1} (x_i - a_i)^2 = -\frac{1}{2\sigma^2} \|\vec{x} - \vec{a}\|^2$$

$$B(\vec{R}) \equiv \sum_{j=0}^{N-1} \ln \left( 1 + \exp \left( B_j(\vec{R}) \right) \right)$$

$$B_j(\vec{R}) \equiv b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} = b_j + \sum_{i=0}^{M-1} x_i W_{i,j} = b_j + \frac{1}{\sigma^2} \vec{x} \cdot \vec{W}_j$$

If the particles are bosons with a hard-core interaction:

$$V_{int}(r_{pq}) = \begin{cases} \infty, & r_{pq} \leq a, \\ 0, & r_{pq} > a, \end{cases}$$

and if the particles are electrons (fermions) with a repulsive Coulomb interaction ( $e = 1$ ):

$$V_{int}(r_{pq}) = \frac{1}{r_{pq}}$$

The behaviors due to these interactions are accounted for in the trial wavefunction by a corresponding Jastrow factor so that the full trial wavefunction is

$$\Psi(\vec{R}) = \frac{1}{Z} \exp \left( A(\vec{R}) + B(\vec{R}) + J(\vec{R}) \right)$$

For bosons with a hard-core interaction:

$$J(\vec{R}) = \sum_{p < q} \ln \left( 1 - \frac{a}{r_{pq}} \right)$$

For electrons with the Coulomb interaction:

$$J(\vec{R}) = \sum_{p < q} \left( \frac{\alpha r_{pq}}{1 + \beta r_{pq}} \right),$$

where  $\beta$  is a variational parameter and  $\alpha$  is 1 when the electrons are anti-parallel and  $\frac{1}{3}$  when the electrons are parallel.

The local energy is defined as

$$E_L = \frac{1}{\Psi} \hat{H} \Psi$$

We want an analytical expression for the local energy in terms of our variational parameters  $\vec{\theta}$ . First, we need the following quantities:

$$\begin{aligned} \nabla_k \{\Psi\} &= \frac{1}{Z} \nabla_k \left\{ \exp(A + B + J) \right\} \\ &= \frac{1}{Z} \exp(A + B + J) \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right) \\ &= \Psi \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right) \end{aligned}$$

$$\begin{aligned} \nabla_k^2 \{\Psi\} &= \nabla_k \left\{ \Psi \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right) \right\} \\ &= \Psi \left( \nabla_k^2 \{A\} + \nabla_k^2 \{B\} + \nabla_k^2 \{J\} \right) + \nabla_k \{\Psi\} \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right) \\ &= \Psi \left( \nabla_k^2 \{A\} + \nabla_k^2 \{B\} + \nabla_k^2 \{J\} + \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right)^2 \right) \end{aligned}$$

$$\frac{1}{\Psi} \nabla_k^2 \{\Psi\} = \nabla_k^2 \{A\} + \nabla_k^2 \{B\} + \nabla_k^2 \{J\} + \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right)^2$$

Let  $\hat{n}_d$ ,  $d = 0, \dots, D-1$ , denote the elementary unit vectors in each of the  $D$  dimensions. Then,

$$\begin{aligned} A(\vec{R}) &\equiv -\frac{1}{2\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} (r_{p,d} - a_{Dp+d})^2, \\ \nabla_k \{A\} &= \sum_{d'=0}^{D-1} \frac{\partial A}{\partial r_{k,d'}} \hat{n}_{d'} = \sum_{d'=0}^{D-1} \frac{\partial}{\partial r_{k,d'}} \left\{ -\frac{1}{2\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} (r_{p,d} - a_{Dp+d})^2 \right\} \hat{n}_{d'} \\ &= \sum_{d'=0}^{D-1} \frac{\partial}{\partial r_{k,d'}} \left\{ -\frac{1}{2\sigma^2} (r_{k,d'} - a_{Dk+d'})^2 \right\} \hat{n}_{d'} = \frac{1}{\sigma^2} \sum_{d=0}^{D-1} (a_{Dk+d} - r_{k,d}) \hat{n}_d \\ \nabla_k^2 \{A\} &= \nabla_k \cdot \nabla_k \{A\} = \nabla_k \cdot \frac{1}{\sigma^2} \sum_{d=0}^{D-1} (a_{Dk+d} - r_{k,d}) \hat{n}_d = \frac{1}{\sigma^2} \sum_{d=0}^{D-1} (-1) = -\frac{D}{\sigma^2} \end{aligned}$$

$$\begin{aligned}
B_j(\vec{R}) &\equiv b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} \\
\nabla_k \{B_j\} &= \sum_{d'=0}^{D-1} \frac{\partial B_j}{\partial r_{k,d'}} \hat{n}_{d'} = \sum_{d'=0}^{D-1} \frac{\partial}{\partial r_{k,d'}} \left\{ b_j + \frac{1}{\sigma^2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} \right\} \hat{n}_{d'} \\
&= \frac{1}{\sigma^2} \sum_{d'=0}^{D-1} \frac{\partial}{\partial r_{k,d'}} [r_{k,d'} W_{Dk+d',j}] \hat{n}_{d'} = \frac{1}{\sigma^2} \sum_{d=0}^{D-1} W_{Dk+d,j} \hat{n}_d \\
\nabla_k^2 \{B_j\} &= 0
\end{aligned}$$

$$\begin{aligned}
B(\vec{R}) &\equiv \sum_{j=0}^{N-1} \ln \left( 1 + \exp \left( B_j(\vec{R}) \right) \right) \\
\nabla_k \{B\} &= \nabla_k \left\{ \sum_{j=0}^{N-1} \ln (1 + \exp (B_j)) \right\} = \sum_{j=0}^{N-1} \nabla_k \left\{ \ln (1 + \exp (B_j)) \right\} \\
&= \sum_{j=0}^{N-1} \frac{1}{1 + \exp(B_j)} \nabla_k \left\{ 1 + \exp(B_j) \right\} = \sum_{j=0}^{N-1} \frac{\exp(B_j)}{1 + \exp(B_j)} \nabla_k \{B_j\} \\
&= \sum_{j=0}^{N-1} \frac{\nabla_k \{B_j\}}{\exp(-B_j) + 1} = \frac{1}{\sigma^2} \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j) + 1} \hat{n}_d \\
\nabla_k^2 \{B\} &= \nabla_k \cdot \nabla_k \{B\} = \nabla_k \cdot \left\{ \sum_{j=0}^{N-1} \frac{\nabla_k \{B_j\}}{\exp(-B_j) + 1} \right\} \\
&= \sum_{j=0}^{N-1} \left[ \nabla_k \left\{ \frac{1}{\exp(-B_j) + 1} \right\} \cdot \nabla_k \{B_j\} + \frac{\nabla_k^2 \{B_j\}}{\exp(-B_j) + 1} \right] \\
&= \sum_{j=0}^{N-1} \frac{\exp(-B_j)}{(\exp(-B_j) + 1)^2} \nabla_k \{B_j\} \cdot \nabla_k \{B_j\} \\
&= \frac{1}{\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j)}{(\exp(-B_j) + 1)^2} \sum_{d=0}^{D-1} W_{Dk+d,j}^2
\end{aligned}$$

$$r_{pq} = \left( \sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})^2 \right)^{1/2}$$

$$\begin{aligned} \nabla_k \{r_{pq}\} &= \frac{1}{2} \left( \sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})^2 \right)^{-1/2} \nabla_k \left\{ \sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})^2 \right\} \\ &= \frac{1}{2r_{pq}} \sum_{d=0}^{D-1} 2(r_{p,d} - r_{q,d})(\delta_{k,p} - \delta_{k,q})\hat{n}_d = \frac{1}{r_{pq}} \sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})(\delta_{k,p} - \delta_{k,q})\hat{n}_d \end{aligned}$$

$$J(\vec{R}) = \sum_{p < q} \ln \left( 1 - \frac{a}{r_{pq}} \right)$$

$$\begin{aligned} \nabla_k \{J\} &= \sum_{p < q} \frac{1}{1 - \frac{a}{r_{pq}}} \nabla_k \left( 1 - \frac{a}{r_{pq}} \right) = \sum_{p < q} \frac{r_{pq}}{r_{pq} - a} \frac{a}{r_{pq}^2} \nabla_k \{r_{pq}\} \\ &= \sum_{p < q} \frac{1}{r_{pq} - a} \frac{a}{r_{pq}} \frac{1}{r_{pq}} \sum_{d=0}^{D-1} (r_{p,d} - r_{q,d})(\delta_{k,p} - \delta_{k,q})\hat{n}_d \\ &= \sum_{d=0}^{D-1} \frac{1}{2} \sum_{p \neq q} \frac{(r_{p,d} - r_{q,d})(\delta_{k,p} - \delta_{k,q})}{r_{pq}^2 \left( \frac{r_{pq}}{a} - 1 \right)} \hat{n}_d \\ &= \sum_{d=0}^{D-1} \frac{1}{2} \left[ \sum_{q \neq k} \frac{r_{k,d} - r_{q,d}}{r_{kq}^2 \left( \frac{r_{kq}}{a} - 1 \right)} - \sum_{k \neq k} \frac{r_{p,d} - r_{k,d}}{r_{pk}^2 \left( \frac{r_{pk}}{a} - 1 \right)} \right] \hat{n}_d \\ &= \sum_{d=0}^{D-1} \frac{1}{2} \left[ 2 \sum_{p \neq k} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \hat{n}_d = \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \hat{n}_d \\ \nabla_k^2 \{J\} &= \sum_{d'=0}^{D-1} \hat{n}_{d'} \frac{\partial}{\partial r_{k,d'}} \cdot \left\{ \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \hat{n}_d \right\} \\ &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{\partial}{\partial r_{k,d}} \left\{ \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right\} \\ &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \left[ (r_{k,d} - r_{p,d}) \frac{\partial}{\partial r_{k,d}} \left\{ r_{kp}^{-2} \left( \frac{r_{kp}}{a} - 1 \right)^{-1} \right\} + \frac{\partial}{\partial r_{k,d}} \{r_{k,d} - r_{p,d}\} \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \\ &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \left[ (r_{k,d} - r_{p,d}) \left[ r_{kp}^{-2} \frac{\partial}{\partial r_{k,d}} \left\{ \left( \frac{r_{kp}}{a} - 1 \right)^{-1} \right\} + \frac{\partial}{\partial r_{k,d}} \{r_{kp}^{-2}\} \left( \frac{r_{kp}}{a} - 1 \right)^{-1} \right] + \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \\ &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \left[ (r_{k,d} - r_{p,d}) \left[ \frac{-1}{ar_{kp}^2} \left( \frac{r_{kp}}{a} - 1 \right)^{-2} \frac{\partial}{\partial r_{k,d}} \{r_{kp}\} + \frac{-2}{r_{kp}^3} \frac{\partial}{\partial r_{k,d}} \{r_{kp}\} \left( \frac{r_{kp}}{a} - 1 \right)^{-1} \right] \right. \\ &\quad \left. + \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \\ &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \left[ -\frac{(r_{k,d} - r_{p,d})}{r_{kp}^2} \left( \frac{r_{kp}}{a} - 1 \right)^{-1} \frac{\partial}{\partial r_{k,d}} \{r_{kp}\} \left[ \frac{1}{a} \left( \frac{r_{kp}}{a} - 1 \right)^{-1} + \frac{2}{r_{kp}} \right] + \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \end{aligned}$$

$$\begin{aligned}
\nabla_k^2\{J\} &= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \left[ -\frac{(r_{k,d} - r_{p,d})}{r_{kp}^2 \left(\frac{r_{kp}}{a} - 1\right)} \left(\frac{r_{k,d} - r_{p,d}}{r_{kp}}\right) \left[\frac{1}{r_{kp} - a} + \frac{2}{r_{kp}}\right] + \frac{1}{r_{kp}^2 \left(\frac{r_{kp}}{a} - 1\right)} \right] \\
&= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{1}{r_{kp}^2 \left(\frac{r_{kp}}{a} - 1\right)} \left[ -\frac{(r_{k,d} - r_{p,d})^2}{r_{kp}} \left[\frac{3r_{kp} - 2a}{r_{kp}(r_{kp} - a)}\right] + 1 \right] \\
&= \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{1}{r_{kp}^2 \left(\frac{r_{kp}}{a} - 1\right)} \left[ \left(\frac{r_{k,d} - r_{p,d}}{r_{kp}}\right)^2 \frac{2a - 3r_{kp}}{r_{kp} - a} + 1 \right]
\end{aligned}$$

Putting all this together...

$$\begin{aligned}
E_L &= \frac{1}{\Psi} \left[ \frac{1}{2} \sum_{p=0}^{P-1} -\nabla_p^2 + \frac{1}{2} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \right] \Psi \\
&= -\frac{1}{2} \sum_{k=0}^{P-1} \frac{1}{\Psi} \nabla_k^2 \Psi + \frac{1}{2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \\
&= -\frac{1}{2} \sum_{k=0}^{P-1} \left( \nabla_k^2 \{A\} + \nabla_k^2 \{B\} + \nabla_k^2 \{J\} + \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right)^2 \right) + \frac{1}{2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \\
&= -\frac{1}{2} \sum_{k=0}^{P-1} \left( -\frac{D}{\sigma^2} + \frac{1}{\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j)}{(\exp(-B_j) + 1)^2} \sum_{d=0}^{D-1} W_{Dk+d,j}^2 \right. \\
&\quad + \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \left[ \left( \frac{r_{k,d} - r_{p,d}}{r_{kp}} \right)^2 \frac{2a - 3r_{kp}}{r_{kp} - a} + 1 \right] \\
&\quad + \left( \frac{1}{\sigma^2} \sum_{d=0}^{D-1} (a_{Dk+d} - r_{k,d}) \hat{n}_d + \frac{1}{\sigma^2} \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j) + 1} \hat{n}_d + \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \hat{n}_d \right)^2 \Bigg) \\
&\quad + \frac{1}{2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \\
&= -\frac{1}{2} \sum_{k=0}^{P-1} \left( \sum_{d=0}^{D-1} \left[ \frac{1}{\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j) W_{Dk+d,j}^2}{(\exp(-B_j) + 1)^2} + \sum_{p=0, p \neq k}^{P-1} \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \left[ \left( \frac{r_{k,d} - r_{p,d}}{r_{kp}} \right)^2 \frac{2a - 3r_{kp}}{r_{kp} - a} + 1 \right] \right] \right. \\
&\quad + \left( \sum_{d=0}^{D-1} \left[ \frac{1}{\sigma^2} (a_{Dk+d} - r_{k,d}) + \frac{1}{\sigma^2} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j) + 1} + \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \hat{n}_d \right)^2 \Bigg) \\
&\quad + \frac{PD}{2\sigma^2} + \frac{1}{2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \\
&= -\frac{1}{2} \sum_{k=0}^{P-1} \sum_{d=0}^{D-1} \left[ \frac{1}{\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j) W_{Dk+d,j}^2}{(\exp(-B_j) + 1)^2} + \sum_{p=0, p \neq k}^{P-1} \frac{1}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \left[ \left( \frac{r_{k,d} - r_{p,d}}{r_{kp}} \right)^2 \frac{2a - 3r_{kp}}{r_{kp} - a} + 1 \right] \right. \\
&\quad + \left( \frac{1}{\sigma^2} (a_{Dk+d} - r_{k,d}) + \frac{1}{\sigma^2} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j) + 1} + \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right)^2 \Bigg] \\
&\quad + \frac{PD}{2\sigma^2} + \frac{1}{2} \sum_{p=0}^{P-1} \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2
\end{aligned}$$

To write  $E_L$  in terms of the inputs  $x_i$ , we define the distance between the  $k$ th and  $p$ th particle as

$$\begin{aligned}
R_i(p) \equiv r_{kp} &= \left( \sum_{d=0}^{D-1} (r_{k,d} - r_{p,d})^2 \right)^{1/2} \\
&= r_{(\text{floor}(i/D))p} = \left( \sum_{d=0}^{D-1} (x_{D(\text{floor}(i/D))+d} - x_{Dp+d})^2 \right)^{1/2}
\end{aligned}$$



Then,

$$\begin{aligned}
E_L &= -\frac{1}{2} \sum_{i=0}^{M-1} \left[ \frac{1}{\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j) W_{i,j}^2}{(\exp(-B_j) + 1)^2} \right. \\
&\quad + a \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{1}{R_i(p)^2 (R_i(p) - a)} \left[ \left( \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)} \right)^2 \frac{2a - 3R_i(p)}{R_i(p) - a} + 1 \right] \\
&\quad + \left( \frac{1}{\sigma^2} (a_i - x_i) + \frac{1}{\sigma^2} \sum_{j=0}^{N-1} \frac{W_{i,j}}{\exp(-B_j) + 1} + a \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)^2 (R_i(p) - a)} \right)^2 \Big] \\
&\quad + \frac{M}{2\sigma^2} + \frac{1}{2} \sum_{i=0}^{M-1} \omega_{i \bmod D}^2 x_i^2 \\
&= -\frac{1}{2\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j)}{(\exp(-B_j) + 1)^2} \sum_{i=0}^{M-1} W_{i,j}^2 \\
&\quad - \frac{a}{2} \sum_{i=0}^{M-1} \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{1}{R_i(p)^2 (R_i(p) - a)} \left[ \left( \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)} \right)^2 \frac{2a - 3R_i(p)}{R_i(p) - a} + 1 \right] \\
&\quad - \frac{1}{2} \sum_{i=0}^{M-1} \left( \frac{1}{\sigma^2} (a_i - x_i) + \frac{1}{\sigma^2} \sum_{j=0}^{N-1} \frac{W_{i,j}}{\exp(-B_j) + 1} + a \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)^2 (R_i(p) - a)} \right)^2 \\
&\quad + \frac{M}{2\sigma^2} + \frac{1}{2} \sum_{i=0}^{M-1} \omega_{i \bmod D}^2 x_i^2 \\
&= -\frac{1}{2\sigma^4} \sum_{j=0}^{N-1} \frac{\exp(-B_j)}{(\exp(-B_j) + 1)^2} \vec{W}_j \cdot \vec{W}_j \\
&\quad - \frac{a}{2} \sum_{i=0}^{M-1} \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{1}{R_i(p)^2 (R_i(p) - a)} \left[ \left( \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)} \right)^2 \frac{2a - 3R_i(p)}{R_i(p) - a} + 1 \right] \\
&\quad - \frac{1}{2} \sum_{i=0}^{M-1} \left( \frac{1}{\sigma^2} (a_i - x_i) + \frac{1}{\sigma^2} \sum_{j=0}^{N-1} \frac{W_{i,j}}{\exp(-B_j) + 1} + a \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)^2 (R_i(p) - a)} \right)^2 \\
&\quad + \frac{M}{2\sigma^2} + \frac{1}{2} (\Omega \% \Omega) \cdot (\vec{x} \% \vec{x})
\end{aligned}$$

In the last line, we have defined a vector  $\vec{\Omega}$  with elements  $\Omega_i = \omega_{i \bmod D}$ . To increase the computational speed, we will precalculate and store the values  $B_j$  in a vector  $\vec{B}$  and store the following factors in a vector  $\vec{f} \in \mathbb{R}^N$ :

$$f_j \equiv \frac{1}{\exp(-B_j) + 1}$$

Then, the local energy becomes

$$\begin{aligned}
E_L = & -\frac{1}{2} \sum_{j=0}^{N-1} \exp(-B_j) \left\| \frac{f_j \vec{W}_j}{\sigma^2} \right\|^2 + \frac{M}{2\sigma^2} + \frac{1}{2} (\Omega \% \Omega) \cdot (\vec{x} \% \vec{x}) \\
& - \frac{a}{2} \sum_{i=0}^{M-1} \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{1}{R_i(p)^2 (R_i(p) - a)} \left[ \left( \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)} \right)^2 \frac{2a - 3R_i(p)}{R_i(p) - a} + 1 \right] \\
& - \frac{1}{2} \sum_{i=0}^{M-1} \left( \frac{1}{\sigma^2} (a_i - x_i) + \frac{1}{\sigma^2} \vec{W}_i^T \vec{f} + a \sum_{p=0, p \neq \text{floor}(i/D)}^{P-1} \frac{x_i - x_{Dp+(i \bmod D)}}{R_i(p)^2 (R_i(p) - a)} \right)^2
\end{aligned}$$

For stochastic gradient descent, we will also need the gradient of the local energy with respect to the RBM parameters  $\vec{\theta} = (\vec{a}, \vec{b}, \hat{W})$ :

$$G_k = \frac{\partial \langle E_L \rangle}{\partial \theta_k} = 2 \left( \left\langle E_L \frac{1}{\Psi} \frac{\partial \Psi}{\partial \theta_k} \right\rangle - \langle E_L \rangle \left\langle \frac{1}{\Psi} \frac{\partial \Psi}{\partial \theta_k} \right\rangle \right)$$

The derivatives of our new wavefunction with respect to the variational parameters are given by:

$$\begin{aligned}
\frac{\partial \Psi}{\partial \vec{a}} &= \Psi \frac{\partial A(\vec{a})}{\partial \vec{a}} = \Psi \frac{\vec{x} - \vec{a}}{\sigma^2} \\
\frac{\partial \Psi}{\partial \vec{b}} &= \Psi \frac{\partial B(\vec{b}, \hat{W})}{\partial \vec{b}} = \Psi \vec{f} \\
\frac{\partial \Psi}{\partial \hat{W}} &= \Psi \frac{\partial B(\vec{b}, \hat{W})}{\partial \hat{W}} = \Psi \frac{\vec{x} \vec{f}^T}{\sigma^2}
\end{aligned}$$

The quantum force on the  $k$ th particle is defined as

$$\vec{F}_k(\vec{R}) = 2 \frac{1}{\Psi} \nabla_k \Psi. \quad (1)$$

Using the derivatives we have already calculated, we have

$$\begin{aligned}
\vec{F}_k(\vec{R}) &= 2 \left( \nabla_k \{A\} + \nabla_k \{B\} + \nabla_k \{J\} \right) \\
&= 2 \left( \frac{1}{\sigma^2} \sum_{d=0}^{D-1} (a_{Dk+d} - r_{k,d}) \hat{n}_d + \frac{1}{\sigma^2} \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j) + 1} \hat{n}_d + \sum_{d=0}^{D-1} \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \hat{n}_d \right) \\
&= 2 \sum_{d=0}^{D-1} \left[ \frac{1}{\sigma^2} (a_{Dk+d} - r_{k,d}) + \frac{1}{\sigma^2} \vec{W}_{Dk+d}^T \vec{f} + \sum_{p=0, p \neq k}^{P-1} \frac{r_{k,d} - r_{p,d}}{r_{kp}^2 \left( \frac{r_{kp}}{a} - 1 \right)} \right] \hat{n}_d
\end{aligned}$$

The Langevin and Fokker-Planck equations give a new position  $y$  from the old position  $x$ :

$$y = x + d\Delta t F(x) + \xi \sqrt{d\Delta t}, \quad (2)$$

where  $d = 0.5$  is the diffusion constant and  $\Delta t \in [0.001, 0.01]$  is a chosen time step.

The transition probability is given by the Green's function

$$G(y, x) = \frac{1}{(4\pi d\Delta t)^{3N/2}} \exp\left(-\frac{(y - x - d\Delta t F(x))^2}{4d\Delta t}\right), \quad (3)$$

so that the Metropolis-Hastings acceptance ratio is

$$A(y, x) = \min\{1, P(y, x)\}, \quad (4)$$

where

$$\begin{aligned} P(y, x) &= \frac{G(x, y)|\Psi_T(y)|^2}{G(y, x)|\Psi_T(x)|^2} \\ &= \exp\left(-\frac{(x - y - d\Delta t F(y))^2}{4d\Delta t}\right) \exp\left(\frac{(y - x - d\Delta t F(x))^2}{4d\Delta t}\right) \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2} \\ &= \exp\left(-\frac{(x - y)^2 - 2(x - y)d\Delta t F(y) + d^2\Delta t^2 F(y)^2}{4d\Delta t}\right) \\ &\quad \times \exp\left(\frac{(y - x)^2 - 2(y - x)d\Delta t F(x) + d^2\Delta t^2 F(x)^2}{4d\Delta t}\right) \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2} \\ &= \exp\left(\frac{2(x - y)d\Delta t(F(y) + F(x)) + d^2\Delta t^2(F(x)^2 - F(y)^2)}{4d\Delta t}\right) \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2} \\ &= \exp\left(\frac{2(x - y)(F(y) + F(x)) + d\Delta t(F(x)^2 - F(y)^2)}{4}\right) \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2} \\ &= \exp\left(\frac{1}{2}(x - y)(F(y) + F(x)) + \frac{1}{4}d\Delta t(F(x)^2 - F(y)^2)\right) \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2} \end{aligned}$$