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 $\hat{W} \in \mathbb{R}^{M \times N}$
- jth column vector of weight matrix $\vec{W}_j \in \mathbb{R}^M$
- variational parameters $\vec{\theta} = (a_0, \cdots, a_{M-1}, b_0, \cdots, b_{N-1}, W_{0,0}, \cdots, W_{M-1,N-1})$

Assume that the input $\vec{x}$ is already mean-centered and scaled so that $\sigma = 1$.

"Energy" of a configuration of nodes:

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

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

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

Hamiltonian for P D-dimensional bosons in a harmonic oscillator:

$$
\hat{H} = \frac{1}{2} \sum_{p=0}^{P-1} \left(-\nabla_p^2 + \sum_{d=0}^{D-1} \omega_d^2 r_{p,d}^2 \right)
$$

This is also the Hamiltonian for non-interacting fermions.
Q: How do we deal with the hard-core interaction?

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

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

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:

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

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

$$\begin{aligned}
\Psi_{\vec{\theta}}(\vec{R}) &= \frac{1}{Z} \exp\left(-\frac{1}{2}\sum_{p=0}^{P-1}\sum_{d=0}^{D-1}(r_{p,d} - a_{Dp+d})^2\right) \prod_{j=0}^{N-1}\left(1 + \exp\left(b_j + \sum_{p=0}^{P-1}\sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j}\right)\right) \\
&= \frac{1}{Z}\exp(A(\vec{R}))B(\vec{R}) \\
A(\vec{R}) &\equiv -\frac{1}{2}\sum_{p=0}^{P-1}\sum_{d=0}^{D-1}(r_{p,d} - a_{Dp+d})^2 = -\frac{1}{2}\sum_{i=0}^{M-1}(x_i - a_i)^2 = -\frac{1}{2}\|\vec{x} - \vec{a}\|^2 \\
B(\vec{R}) &\equiv \prod_{j=0}^{N-1}\left(1 + \exp\left(B_j(\vec{R})\right)\right) \\
B_j(\vec{R}) &\equiv b_j + \sum_{p=0}^{P-1}\sum_{d=0}^{D-1} r_{p,d} W_{Dp+d,j} = b_j + \sum_{i=0}^{M-1} x_i W_{i,j} = b_j + \vec{x} \cdot \vec{W}_j
\end{aligned}$$

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:

$$
\begin{aligned}
\nabla_k \Psi_{\vec{\theta}}(\vec{R}) &= \frac{1}{Z}\nabla_k \left[\exp(A)B\right] = \frac{1}{Z}\left(\exp(A)\nabla_k[B] + \nabla_k[\exp(A)]B\right) \\
&= \frac{1}{Z}\exp(A)\left(\nabla_k[B] + \nabla_k[A]B\right)
\end{aligned}
$$

$$
\begin{aligned}
\nabla_k^2 \Psi_{\vec{\theta}}(\vec{R}) &= \frac{1}{Z}\nabla_k \left[\exp(A)\left(\nabla_k[B] + \nabla_k[A]B\right)\right] \\
&= \frac{1}{Z}\left(\exp(A)\nabla_k\left[\nabla_k[B] + \nabla_k[A]B\right] + \nabla_k[\exp(A)]\left(\nabla_k[B] + \nabla_k[A]B\right)\right) \\
&= \frac{1}{Z}\Big(\exp(A)\left(\nabla_k^2[B] + \nabla_k[A]\nabla_k[B] + \nabla_k^2[A]B\right) + \exp(A)\nabla_k[A]\left(\nabla_k[B] + \nabla_k[A]B\right)\Big) \\
&= \frac{1}{Z}\exp(A)\left(\nabla_k^2[B] + \nabla_k[A]\nabla_k[B] + \nabla_k^2[A]B + \nabla_k[A]\nabla_k[B] + \left(\nabla_k[A]\right)^2 B\right) \\
&= \frac{1}{Z}\exp(A)\left(\nabla_k^2[B] + 2\nabla_k[A]\nabla_k[B] + \nabla_k^2[A]B + \left(\nabla_k[A]\right)^2 B\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\Psi_{\vec{\theta}}(\vec{R})}\nabla_k^2 \Psi_{\vec{\theta}}(\vec{R}) &= \frac{Z\exp(-A)}{B}\frac{1}{Z}\exp(A)\left(\nabla_k^2[B] + 2\nabla_k[A]\nabla_k[B] + \nabla_k^2[A]B + \left(\nabla_k[A]\right)^2 B\right) \\
&= \frac{\nabla_k^2[B]}{B} + 2\frac{\nabla_k[A]\nabla_k[B]}{B} + \nabla_k^2[A] + \left(\nabla_k[A]\right)^2
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
\nabla_k^2[B] &= \nabla_k \cdot \nabla_k[B] \\
&= \left(\sum_{d'=0}^{D-1} \hat{n_{d'}} \frac{\partial}{\partial r_{k,d'}}\right) \cdot \left[\sum_{d=0}^{D-1} \left(\sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \prod_{j'\neq j} (1+\exp(B_{j'}))\right) \hat{n_d}\right] \\
&= \sum_{d=0}^{D-1} \frac{\partial}{\partial r_{k,d}} \left[\sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \prod_{j'\neq j} (1+\exp(B_{j'}))\right] \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \frac{\partial}{\partial r_{k,d}} \left[\exp(B_j) \prod_{j'\neq j} (1+\exp(B_{j'}))\right] \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \left(\exp(B_j) \frac{\partial}{\partial r_{k,d}} \left[\prod_{j'\neq j} (1+\exp(B_{j'}))\right] + \frac{\partial}{\partial r_{k,d}}[\exp(B_j)] \prod_{j'\neq j} (1+\exp(B_{j'}))\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \left(\exp(B_j) \sum_{j'=0,\ j'\neq j}^{N-1} \frac{\partial}{\partial r_{k,d}}[1+\exp(B_{j'})] \prod_{j''\neq j,j'} (1+\exp(B_{j''})) \right. \\
&\qquad \left. + \exp(B_j) \frac{\partial}{\partial r_{k,d}}[B_j] \prod_{j'\neq j} (1+\exp(B_{j'}))\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \left(\sum_{j'=0,\ j'\neq j}^{N-1} \exp(B_{j'}) \frac{\partial}{\partial r_{k,d}}[B_{j'}] \prod_{j''\neq j,j'} (1+\exp(B_{j''})) \right. \\
&\qquad \left. + \frac{\partial}{\partial r_{k,d}}[B_j] \prod_{j'\neq j} (1+\exp(B_{j'}))\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \left(\sum_{j'=0,\ j'\neq j}^{N-1} \exp(B_{j'}) W_{Dk+d,j'} \prod_{j''\neq j,j'} (1+\exp(B_{j''})) \right. \\
&\qquad \left. + W_{Dk+d,j} \prod_{j'\neq j} (1+\exp(B_{j'}))\right) \\
\frac{\nabla_k^2[B]}{B} &= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \left(\sum_{j'=0,\ j'\neq j}^{N-1} \exp(B_{j'}) W_{Dk+d,j'} \frac{\prod_{j''\neq j,j'} (1+\exp(B_{j''}))}{\prod_{j''=0}^{N-1} (1+\exp(B_{j''}))} \right. \\
&\qquad \left. + W_{Dk+d,j} \frac{\prod_{j'\neq j} (1+\exp(B_{j'}))}{\prod_{j'=0}^{N-1} (1+\exp(B_{j'}))}\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \exp(B_j) \left(\sum_{j'=0,\ j'\neq j}^{N-1} \exp(B_{j'}) W_{Dk+d,j'} \frac{1}{(1+\exp(B_j))(1+\exp(B_{j'}))} \right. \\
&\qquad \left. + W_{Dk+d,j} \frac{1}{1+\exp(B_j)}\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} W_{Dk+d,j} \frac{\exp(B_j)}{1+\exp(B_j)} \left(\sum_{j'=0,\ j'\neq j}^{N-1} W_{Dk+d,j'} \frac{\exp(B_{j'})}{1+\exp(B_{j'})} + W_{Dk+d,j}\right) \\
&= \sum_{d=0}^{D-1} \sum_{j=0}^{N-1} \frac{W_{Dk+d,j}}{\exp(-B_j)+1} \left(\sum_{j'=0,\ j'\neq j}^{N-1} \frac{W_{Dk+d,j'}}{\exp(-B_{j'})+1} + W_{Dk+d,j}\right)
\end{aligned}
$$

Putting all this together...

$$
\begin{aligned}
E_L &= \frac{1}{\Psi_{\vec{\theta}}(\vec{R})}\left[\frac{1}{2}\sum_{p=0}^{P-1}\left(-\nabla_p^2+\sum_{d=0}^{D-1}\omega_d^2 r_{p,d}^2\right)\right]\Psi_{\vec{\theta}}(\vec{R}) \\
&= \sum_{p=0}^{P-1}\left[-\frac{1}{2}\left(\frac{\nabla_p^2[B]}{B}+2\frac{\nabla_p[A]\nabla_p[B]}{B}+\nabla_p^2[A]+(\nabla_p[A])^2\right)+\sum_{d=0}^{D-1}\frac{1}{2}\omega_d^2 r_{p,d}^2\right] \\
&= \sum_{p=0}^{P-1}\left[-\frac{1}{2}\frac{\nabla_p^2[B]}{B}-\frac{\nabla_p[A]\nabla_p[B]}{B}-\frac{1}{2}\nabla_p^2[A]-\frac{1}{2}\left(\nabla_p[A]\right)^2+\sum_{d=0}^{D-1}\frac{1}{2}\omega_d^2 r_{p,d}^2\right] \\
&= \sum_{p=0}^{P-1}\Bigg[-\frac{1}{2}\sum_{d=0}^{D-1}\sum_{j=0}^{N-1}\frac{W_{Dp+d,j}}{\exp(-B_j)+1}\left(\sum_{j'=0,\ j'\neq j}^{N-1}\frac{W_{Dp+d,j'}}{\exp(-B_{j'})+1}+W_{Dp+d,j}\right) \\
&\qquad -\left(\sum_{d'=0}^{D-1}(a_{Dp+d'}-r_{p,d'})\hat{n_{d'}}\right)\cdot\left(\sum_{d=0}^{D-1}\sum_{j=0}^{N-1}\frac{W_{Dp+d,j}}{\exp(-B_j)+1}\hat{n_d}\right) \\
&\qquad -\frac{1}{2}(-D)-\frac{1}{2}\sum_{d=0}^{D-1}(a_{Dp+d}-r_{p,d})^2+\sum_{d=0}^{D-1}\frac{1}{2}\omega_d^2 r_{p,d}^2\Bigg] \\
&= \sum_{p=0}^{P-1}\Bigg[-\frac{1}{2}\sum_{d=0}^{D-1}\sum_{j=0}^{N-1}\frac{W_{Dp+d,j}}{\exp(-B_j)+1}\left(\sum_{j'=0,\ j'\neq j}^{N-1}\frac{W_{Dp+d,j'}}{\exp(-B_{j'})+1}+W_{Dp+d,j}\right) \\
&\qquad -\sum_{d=0}^{D-1}\sum_{j=0}^{N-1}\frac{(a_{Dp+d}-r_{p,d})W_{Dp+d,j}}{\exp(-B_j)+1}+\frac{D}{2}-\frac{1}{2}\sum_{d=0}^{D-1}(a_{Dp+d}-r_{p,d})^2+\sum_{d=0}^{D-1}\frac{1}{2}\omega_d^2 r_{p,d}^2\Bigg] \\
&= \frac{PD}{2}+\sum_{p=0}^{P-1}\sum_{d=0}^{D-1}\Bigg[-\frac{1}{2}\sum_{j=0}^{N-1}\sum_{j'=0,\ j'\neq j}^{N-1}\frac{W_{Dp+d,j}W_{Dp+d,j'}}{(\exp(-B_j)+1)(\exp(-B_{j'})+1)}-\frac{1}{2}\sum_{j=0}^{N-1}\frac{W_{Dp+d,j}^2}{\exp(-B_j)+1} \\
&\qquad -\sum_{j=0}^{N-1}\frac{(a_{Dp+d}-r_{p,d})W_{Dp+d,j}}{\exp(-B_j)+1}-\frac{1}{2}(a_{Dp+d}-r_{p,d})^2+\frac{1}{2}\omega_d^2 r_{p,d}^2\Bigg]
\end{aligned}
$$

In terms of the inputs x_i, $i = 0, ..., M-1$,

$$
\begin{aligned}
E_L = \frac{M}{2} &+ \sum_{i=0}^{M-1} \left[-\sum_{j=0}^{N-1} \sum_{j'<j} \frac{W_{i,j} W_{i,j'}}{(\exp(-B_j)+1)(\exp(-B_{j'})+1)} - \frac{1}{2} \sum_{j=0}^{N-1} \frac{W_{i,j}^2}{\exp(-B_j)+1} \right. \\
&\qquad \left. - \sum_{j=0}^{N-1} \frac{(a_i - x_i) W_{i,j}}{\exp(-B_j)+1} - \frac{1}{2}(a_i - x_i)^2 + \frac{1}{2} \omega_{i\text{mod}D}^2 x_i^2 \right] \\
= \frac{M}{2} &+ \sum_{j=0}^{N-1} \left[-\sum_{j'<j} \sum_{i=0}^{M-1} \frac{W_{i,j} W_{i,j'}}{(\exp(-B_j)+1)(\exp(-B_{j'})+1)} - \frac{1}{2} \sum_{i=0}^{M-1} \frac{W_{i,j}^2}{\exp(-B_j)+1} - \sum_{i=0}^{M-1} \frac{(a_i - x_i) W_{i,j}}{\exp(-B_j)+1} \right] \\
&+ \sum_{i=0}^{M-1} \left[-\frac{1}{2}(a_i - x_i)^2 + \frac{1}{2} \omega_{i\text{mod}D}^2 x_i^2 \right] \\
= \frac{M}{2} &+ \sum_{j=0}^{N-1} \left[-\sum_{j'<j} \frac{\vec{W}_j \cdot \vec{W}_{j'}}{(\exp(-B_j)+1)(\exp(-B_{j'})+1)} - \frac{1}{2} \frac{\vec{W}_j \cdot \vec{W}_j}{\exp(-B_j)+1} - \frac{(\vec{a} - \vec{x}) \cdot \vec{W}_j}{\exp(-B_j)+1} \right] \\
&- \frac{1}{2}(\vec{a} - \vec{x}) \cdot (\vec{a} - \vec{x}) + \frac{1}{2} \sum_{i=0}^{M-1} \omega_{i\text{mod}D}^2 x_i^2 \\
= \frac{M}{2} &- \sum_{j=0}^{N-1} \left[\frac{\vec{W}_j}{\exp(-B_j)+1} \cdot \left(\sum_{j'<j} \frac{\vec{W}_{j'}}{\exp(-B_{j'})+1} + \frac{1}{2} \vec{W}_j + (\vec{a} - \vec{x}) \right) \right] - \frac{1}{2} \|\vec{a} - \vec{x}\|^2 + \frac{1}{2} (\Omega \% \Omega) \cdot (\vec{x} \% \vec{x})
\end{aligned}
$$

In the last line, we have defined a vector $\vec{\Omega}$ with elements $\Omega_i = \omega_{i\text{mod}D}$. For more efficient computation of E_L I will store the factors $f_j \equiv \frac{1}{\exp(-B_j)+1}$ in a vector $\vec{f}$ so that

$$
E_L = \frac{M}{2} - \sum_{j=0}^{N-1} \left[f_j \vec{W}_j \cdot \left(\sum_{j'<j} f_{j'} \vec{W}_{j'} + \frac{1}{2} \vec{W}_j + (\vec{a} - \vec{x}) \right) \right] - \frac{1}{2} \|\vec{a} - \vec{x}\|^2 + \frac{1}{2} (\Omega \% \Omega) \cdot (\vec{x} \% \vec{x})
$$

For stochastic gradient descent, we will also need the gradient of the local energy with respect to the RBM parameters $\vec{\theta}$:

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

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

$$
\begin{aligned}
\frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial a_k} &= \Psi_{\vec{\theta}} (x_k - a_k) & \longrightarrow \frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial \vec{a}} &= \Psi_{\vec{\theta}} (\vec{x} - \vec{a}) \\
\frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial b_k} &= \Psi_{\vec{\theta}} \frac{1}{\exp(-B_k)+1} & \longrightarrow \frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial \vec{b}} &= \Psi_{\vec{\theta}} \vec{f} \\
\frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial W_{k,l}} &= \Psi_{\vec{\theta}} \frac{x_k}{\exp(-B_l)+1} & \longrightarrow \frac{\partial \Psi_{\vec{\theta}}(\vec{R})}{\partial \vec{W}_j} &= \Psi_{\vec{\theta}} f_j \vec{x}
\end{aligned}
$$