To find the local energy minimum of the NQS wavefunction the gradient of the local energy with respect to the variational parameter $\alpha = \{\mathbf{a}, \mathbf{b}, \mathbf{W}\}$ needs to be computed. For each variational parameter $\alpha_i \in \alpha$ the gradient is defined as in equation 1

$$G_i = \frac{\partial \langle E_L \rangle}{\partial \alpha_i} = 2\left(\langle E_L \frac{1}{\Psi} \frac{\partial \Psi}{\partial \alpha_i} \rangle - \langle E_L \rangle \langle \frac{1}{\Psi} \frac{\partial \Psi}{\partial \alpha_i} \rangle \right) \tag{1}$$

While the NQS is defined for the Gaussian Binary RBM as

$$\Psi(\mathbf{X}) = \frac{1}{Z} \exp\left(-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2} \right) \prod_j^N \left(1 + \exp\left(b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2} \right) \right) \tag{2}$$

We now find it useful to use the identity

$$\frac{d}{dx} \ln f(x) = \frac{1}{f(x)} \frac{d}{dx} f(x) \tag{3}$$

Taking the logarithm of the NQS wavefunction then gives

$$\ln(\Psi(\mathbf{X})) = \ln \frac{1}{Z} - \sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2} + \sum_j^N \ln \left[1 + \exp\left(b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2} \right) \right] \tag{4}$$

The final result for the gradient of the vector and matrix components of α_i are then

$$\frac{\partial \ln(\Psi(\mathbf{X}))}{\partial a_k} = \frac{1}{\sigma^2}(X_k - a_k) \tag{5}$$

$$\frac{\partial \ln(\Psi(\mathbf{X}))}{\partial b_k} = \left(1 + \exp\left(-b_k - \sum_i^M \frac{X_i w_{ik}}{\sigma^2} \right) \right)^{-1} \tag{6}$$

$$\frac{\partial \ln(\Psi(\mathbf{X}))}{\partial w_{kn}} = \frac{1}{\sigma^2} \left(1 + \exp\left(-b_k - \sum_i^M \frac{X_i w_{ik}}{\sigma^2} \right) \right)^{-1} \tag{7}$$