Computational Tools for Quantum Many-Body Physics (WS 24/25)
Exercise Sheet 6

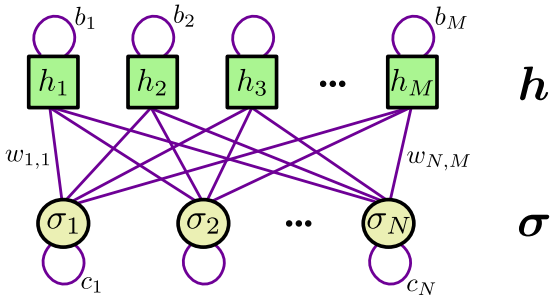
Jan Budich

Online: Thu, 23 Jan 2025
Discussion: Thu, 06 Feb 2025

The purpose of this exercise sheet is to write and benchmark a basic variational Monte Carlo (VMC) code for calculating the ground-state of a QMB system, here exemplified by the one-dimensional spin 1/2 XXZ model known from previous problem sheets, which is described by the Hamiltonian

$$H = -J \sum_j \left(S^x_j S^x_{j+1} + S^y_j S^y_{j+1} \right) + \Delta \sum_j S^z_j S^z_{j+1} \ . \tag{1}$$

The variational ansatz $|\psi_{\boldsymbol{w}}\rangle$ used in this sheet is known as a restricted Boltzmann machine (RBM) state [1], i.e. derived from an artificial neural network. The RBM state amplitude $\psi_{\boldsymbol{w}}(\boldsymbol{\sigma}) = \langle \boldsymbol{\sigma} | \psi_{\boldsymbol{w}} \rangle$ for a given spin configuration $\boldsymbol{\sigma} = \{\sigma_1, ..., \sigma_N\}$ is obtained from the network illustrated below, where the yellow circles correspond to the physical spins, the green boxes are hidden (or auxiliary) units taken to be classical spins with values $h_j = \pm 1$, and the purple links are the network weights that correspond to the (complex) variational parameters to be optimized in the VMC computation. The network connectivity is encoded in the network-energy function $\mathcal{E}_{\rm nw}(\boldsymbol{\sigma}, \boldsymbol{h}; \boldsymbol{w})$ defined below.



$$\mathcal{E}_{\rm nw}(\boldsymbol{\sigma}, \boldsymbol{h}; \boldsymbol{w}) = -\sum_{i=1}^{N} c_i \sigma_i + \sum_{j=1}^{M} \Big(b_j + \sum_{i=1}^{N} \sigma_i w_{i,j} \Big) h_j$$

To obtain the RBM state amplitude $\langle \boldsymbol{\sigma} | \psi_{\boldsymbol{w}} \rangle = \psi_{\boldsymbol{w}}(\boldsymbol{\sigma})$, depending on the spin configuration $\boldsymbol{\sigma}$ and the set of variational parameters $\boldsymbol{w} = \{c_i, b_j, w_{i,j}\}$, one eliminates the dependence on the hidden units h_j by summing over all their possible configurations the Boltzmann-like factor $\mathrm{e}^{-\mathcal{E}_{\rm nw}(\boldsymbol{\sigma}, \boldsymbol{h}; \boldsymbol{w})}$. This sum amounts to generating correlations among the physical spins, encoded in the following functional form of the variational ansatz

$$\psi_{\boldsymbol{w}}(\boldsymbol{\sigma}) = \sum_{\{\boldsymbol{h}\}} \mathrm{e}^{-\mathcal{E}_{\rm nw}(\boldsymbol{\sigma}, \boldsymbol{h}; \boldsymbol{w})} = \mathrm{e}^{\sum_{i=1}^{N} c_i \sigma_i} \prod_{j=1}^{M} 2 \cosh \left(b_j + \sum_{i=1}^{N} \sigma_i w_{i,j} \right) \ . \tag{2}$$

It is worth pointing out that $\psi_{\boldsymbol{w}}(\boldsymbol{\sigma})$ is the *non-normalized* state amplitude, as the norm of the state is generally (i.e. for large systems) not accessible.

Remark: The model for this exercise sheet is chosen such that you can benchmark your results against the ED and DMRG results obtained in the previous exercises. However, it is important to mention that, unlike DMRG, VMC does not suffer from major limitations in going to higher spatial dimensions.

1. **Calculating the local energy estimator**
   The first quantity that needs to be calculated (stochastically estimated) is the energy expectation value $E_{\boldsymbol{w}} = \frac{\langle \psi_{\boldsymbol{w}} | H | \psi_{\boldsymbol{w}} \rangle}{\langle \psi_{\boldsymbol{w}} | \psi_{\boldsymbol{w}} \rangle}$ over the variational RBM state $|\psi_{\boldsymbol{w}}\rangle$. It can be expressed as

$$E_{\boldsymbol{w}} = \sum_{\boldsymbol{\sigma}} \frac{|\langle \boldsymbol{\sigma} | \psi_{\boldsymbol{w}} \rangle|^2}{\langle \psi_{\boldsymbol{w}} | \psi_{\boldsymbol{w}} \rangle} E_{\rm loc}(\boldsymbol{\sigma}) \ . \tag{3}$$

which can be estimated by sampling the probability distribution $P_{\boldsymbol{w}}(\boldsymbol{\sigma}) \equiv \frac{|\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle|^2}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}$, where $E_{\rm loc}(\boldsymbol{\sigma})$ is the local energy estimator

$$E_{\rm loc}(\boldsymbol{\sigma}) = \frac{\langle \boldsymbol{\sigma}|H|\psi_{\boldsymbol{w}}\rangle}{\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle} \ . \tag{4}$$

Write a routine that calculates $E_{\rm loc}(\boldsymbol{\sigma})$ for a given spin configuration $\boldsymbol{\sigma}$ and a given set of variational parameters $\boldsymbol{w}$. Notice that for calculating $E_{\rm loc}(\boldsymbol{\sigma})$ one only needs to know the action of H on $|\boldsymbol{\sigma}\rangle$, i.e. one loop over the number N of sites is sufficient for its calculation (and *not* a loop over the whole Hilbert space).

2. **Calculating the gradients in parameter space**
Since the goal of VMC is to optimize the variational parameters $\boldsymbol{w}$ so as to find minimum of the energy $E_{\boldsymbol{w}}$, the second quantity that needs to be estimated is $\nabla_{\boldsymbol{w}} E_{\boldsymbol{w}}$, i.e. derivatives of the energy expectation value with respect to the variational parameters. Give an expression for the logarithmic derivatives of the RBM state amplitude $\psi_{\boldsymbol{w}}(\boldsymbol{\sigma})$ with respect to the real and imaginary part of the variational parameters

$$O_{{\rm re},k}(\boldsymbol{\sigma}) \equiv \frac{\partial \log \psi_{\boldsymbol{w}}(\boldsymbol{\sigma})}{\partial w_k^{\rm re}} \ , \quad O_{{\rm im},k}(\boldsymbol{\sigma}) \equiv \frac{\partial \log \psi_{\boldsymbol{w}}(\boldsymbol{\sigma})}{\partial w_k^{\rm im}} \tag{5}$$

where $w_k^{\rm re/im}$ is the real/imaginary part of the k-th variational parameter in the set $\boldsymbol{w}$. Show that

$$\frac{\langle \partial_{w_k^{\rm re/im}} \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle} = \sum_{\boldsymbol{\sigma}} \frac{|\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle|^2}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle} O^*_{{\rm re/im},k}(\boldsymbol{\sigma}) \tag{6}$$

holds, which also can be estimated by sampling from $P_{\boldsymbol{w}}(\boldsymbol{\sigma}) \equiv \frac{|\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle|^2}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}$. Give an expression for the derivatives $\nabla_{\boldsymbol{w}} E_{\boldsymbol{w}}$ of energy expectation value $E_{\boldsymbol{w}}$ of Eq. (3) with respect to the real and imaginary part of the variational parameters in terms of the logarithmic derivatives $O_{{\rm re/im},k}(\boldsymbol{\sigma})$.

3. **Metropolis sampling**
Implement a routine that uses the Metropolis algorithm to sample from the probability distribution $P_{\boldsymbol{w}}(\boldsymbol{\sigma}) \equiv \frac{|\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle|^2}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}$. Keeping in mind that the expectation value $\langle A\rangle = \frac{\langle \psi_{\boldsymbol{w}}|A|\psi_{\boldsymbol{w}}\rangle}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}$ over the variational state $|\psi_{\boldsymbol{w}}\rangle$ can be expressed by Monte Carlo average over the N_s samples $\boldsymbol{\sigma}_n$ of $P_{\boldsymbol{w}}(\boldsymbol{\sigma})$

$$\langle A\rangle = \sum_{\boldsymbol{\sigma}} P_{\boldsymbol{w}}(\boldsymbol{\sigma})\, A_{\rm loc}(\boldsymbol{\sigma}) \simeq \frac{1}{N_s} \sum_{n=1}^{N_s} A_{\rm loc}(\boldsymbol{\sigma}_n) \quad \text{with} \quad A_{\rm loc}(\boldsymbol{\sigma}) = \frac{\langle \boldsymbol{\sigma}|A|\psi_{\boldsymbol{w}}\rangle}{\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle} \tag{7}$$

calculate the values of $E_{\boldsymbol{w}}$ and $\nabla_{\boldsymbol{w}} E_{\boldsymbol{w}}$ from the samples.

Hint: Remember that the XXZ model (1) conserves the total z magnetization. If the magnetization sector to which the ground-state of H belongs is known, it is therefore sufficient to sample only spin configurations having that fixed value of z magnetization.

4. **Iterative optimization of the variational parameters**
The next step is to put together a full working VMC code. Use $J = 1$ and different values of Δ, e.g. $\Delta = 0, 1$ and 2, and $N = 16$ so that you can benchmark your code against the results for the ground-state energy obtained by ED. Choose the number of hidden units M to be $M = \alpha N$ with $\alpha = 1, 2, 3$ and compare the accuracy of these three cases. The general procedure is as follows:

   (a) Start from a random configuration of variational parameters $\boldsymbol{w}$.

   (b) Generate N_s samples of the probability distribution $P_{\boldsymbol{w}}(\boldsymbol{\sigma}) \equiv \frac{|\langle \boldsymbol{\sigma}|\psi_{\boldsymbol{w}}\rangle|^2}{\langle \psi_{\boldsymbol{w}}|\psi_{\boldsymbol{w}}\rangle}$.

   (c) Estimate $E_{\boldsymbol{w}}$ and $\nabla_{\boldsymbol{w}} E_{\boldsymbol{w}}$ from the samples.

   (d) Update the variational parameters $\boldsymbol{w}$ by descending in the energy landscape, i.e. the update vector $d\boldsymbol{w}$ reads as

   $$d\boldsymbol{w} = -\gamma\, \nabla_{\boldsymbol{w}} E_{\boldsymbol{w}}, \tag{8}$$

   where γ is called *learning rate* and controls the length of the step made in the parameter space.

   (e) Repeat from (b) until the energy $E_{\boldsymbol{w}}$ has converged to a minimum.

Try to increase the size of the system to $N \geq 30$ (you can still compare your results to exact results in the case $\Delta = 0$ with open boundaries, as the XXZ model maps to free spinless fermions by means of Jordan-Wigner transformation).

Hint: You can make the learning rate γ "time"-dependent, in the sense that you can decrease gradually its value during the iterations, so as to make the search of the minimum more accurate towards the end of the optimization. In addition, you can try to modify the optimization step of Eq. (8) by implementing a AdaGrad-like method [2].

References

[1] G. Carleo, M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science **355**, 602 (2017).

[2] J. Duchi, E. Hazan, Y. Singer, *Adaptive subgradient methods for online learning and stochastic optimization*, Journal of Machine Learning Research **12**, 2121 (2011).