1 Unitary Operations and RBMs

The state of an N-spin system is described by a wavefunction $\Psi(s)$ with $s \in \{0,1\}^N$. $\Psi(s)$ can be encoded in a so-called Boltzmann Machine, a simple two-layer ANN. It consists of N_v visible and N_h hidden nodes and is specified by parameters $a \in \mathbb{R}^{N_v}, b \in \mathbb{R}^{N_h}$ and $w \in \mathbb{R}^{N_h \times N_v}$. The energy function

$$E(v,h) = v^T a + h^T b + h^T w v \tag{1}$$

allows one to express probability distributions defined on $\mathbb{R}^{N_v}$ according to

$$P(v) = \frac{\sum_h e^{-E(v,h)}}{\sum_v \sum_h e^{-E(v,h)}}. \tag{2}$$

Similarly, if we allow complex $a \in \mathbb{C}^{N_v}, b \in \mathbb{C}^{N_h}$ and $w \in \mathbb{C}^{N_h \times N_v}$ we can map each spin configuration $s \in \{0,1\}^{N_S}$ of system with N_S spins to a complex amplitude. This is then the wavefunction

$$\Psi(s) = \frac{\sum_h e^{-E(s,h)}}{\sum_s \sum_h e^{-E(s,h)}}. \tag{3}$$

Consider now a 1-body unitary operator $O \in U(2)$. It is completely specified by its four matrix elements $O_{ss'}$ where $s, s' \in \{0,1\}$. Equivalently, we can express this Operator in terms of an exponential function:

$$O(s,s') = A \exp\left(\alpha s + \beta s' + \omega s s'\right). \tag{4}$$

To do this we simply associate

$$\begin{aligned} A &= O_{00} \\ \alpha &= \ln\left(\frac{O_{10}}{A}\right) \\ \beta &= \ln\left(\frac{O_{01}}{A}\right) \\ \omega &= \ln\left(\frac{O_{11}}{A}\right) - (\alpha + \beta). \end{aligned}$$

Clearly the function $O(s,s')$ defined this way behaves like the operator. The wavefunction $\Psi'(s)$ after O has acted upon the initial state $Psi(s)$ is given by

$$\begin{aligned} \Psi'(s) &= \sum_{s'} O_{ss'} \Psi(s_1, \ldots, s', \ldots, s_N) \\ &= \sum_{s'} \exp\left(\alpha s + \beta s' + \omega s s'\right) \Psi(s_1, \ldots, s', \ldots, s_N) \end{aligned} \tag{5}$$

Expressing the RHS wave-function now in terms of our RBM (cf. 3), the sum over spin s' can be associated with another hidden node.

2 Unitary Operations and POVM Distributions

Suppose now the RBM encodes some probability distribution $P(\vec{a})$ of an informationally complete POVM measurement. This distribution then uniquely specifies some quantum state. Let $O_{\vec{a},\vec{b}}$ now denote a single-body unitary operator acting on POVM distribution vectors. Is it still possible to rewrite this operator as an exponential function and describe the effect of the unitary operation as the insertion of an additional node into the RBM?

Let us first consider the operator. The principal goal is to write it in exponential form such that it can be absorbed into the energy functional of a slightly modified architecture. There exist several ways of implementing POVMs outcoms in an RBM. For reasons which will become apparent later we will choose binary encoding. This means each outcome is encoded by pairs of two nodes in the network. Binary nodes then lead to the configurations in 1 A single-qubit

POVM outcome	node configuration
1	(0,0)
2	(0,1)
3	(1,0)
4	(1,1)

Figure 1: 4-POVM outcomes implemented by binary nodes.

gate in the POVM formalism corresponds to a complex (4×4)-matrix. Its 16 entries need to be represented in their entirety in the exponential representation of this operator. Rows in this matrix can be associated with a fixed "target outcome" while entries in the same column correspond to the same "initial outcome". Let us denote the two nodes encoding the target outcome by (v_1, v_2) and the nodes encoding the initial outcome by (h_1, h_2). The entry O_{ij} corresponds to the transition element for $j \to i$, where $i, j \in \{1, \ldots, 4\}$ represent POVM measurements. These states can also be described in terms of the nodes that encode them. Then we have transitions $(h_1^j, h_2^j) \to (v_1^i, v_2^i)$. Each transition (to which one specific matrix element is associated) can be represented by a unique 4-tuple $(h_1^j, h_2^j, v_1^i, v_2^i)$ where each entry takes binary values.

Consider now products of these nodes. All possible products are given by

$$\{h_1^i h_2^j v_1^k v_2^l \;:\; (i,j,k,l) \in \{0,1\}^4\} \tag{6}$$

I will argue that the value of these variable-products can be used to identify specific transitions. It is clear that $h_1 h_2 v_1 v_2$ only evaluates to 1 if all variables are equal to one. Hence this outcome can be used to identify the $(4 \to 4)$ transition. We hence include this product in the exponent and multiply it with a coefficient $\ln(O_{44})$ including some correction factors which we will discuss later. The energy functional we choose is

$$E(v_1, v_2, h_1, h_2) = \sum_{i=0,1} \sum_{j=0,1} \sum_{k=0,1} \sum_{l=0,1} c_{ijkl} h_1^i h_2^j v_1^k v_2^l \tag{7}$$

We recognize that the transition $1 \to 1$ encoded by binary variable products in the energy functional is just c_{0000} and always evaluates to 1, irrespective of the specific input (v_1, v_2, h_1, h_2). We thus write $c_{0000} = \ln(O_{11})$. For input $(0, 0, 0, 0)$ all other terms don't contribute since at least one factor is always zero. When looking at configurations that involve some nodes not equal to zero, we have to make the following consideration: Say we look at the transition $4 \to 1$. This corresponds to $v_1 = 0, v_2 = 0 h_1 = 1, h_2 = 2$. All products unequal to zero are then $h1, h2, h1h2$. We cannot simply identify $c_{0011} = O_{14}$, but instead have to account for the fact that we have "unwanted" activation products ($h1, h2$ corresponding to the transitions $2 \to 1, 3 \to 1$). Thus we exclude these by an appropriate coefficient choice: $c_{0011} = O_{14} - c_{0010} - c_{0001}$. This procedure can be applied to every c_{ijkl} and we end up with a unique relation $c_{ijkl} = f_{ijkl}(O)$.

Since we showed that the POVM-Operator can be represented as an exponential function, the occuring sum over "prior" outcomes b_i can be – together with the operator in exponent form – associated with additional hidden nodes added to the network. For our one-body operation, this leads to already two (rather than 1 for the case of spin operators) additional hidden nodes. These interact with other hidden nodes and hence we have an Unrestricted Boltzmann machine. This makes information extraction complicated, since no analytical expression for the probability encoded by the UBM can be found anymore due to the interaction of the hidden nodes. What can be done however is to transform the unrestricted Boltzmann machine into a Deep Boltzmann Machine by the addition of a second hidden layer that includes the two nodes added by the one-body operation. What were formally interactions within the hidden layer are now interaction between the first and second hidden layer. The two additional nodes however also interact with the visible units. As we do not allow direct interaction between the visible and second hidden layer in the DBM, additional nodes can be inserted in the first hidden layer that mediate the interaction between the visible and second hidden layer.

Then information from the DBM could be extracted as follows:

1. Consider hidden layer 1 and hidden layer 2 as a simple RBM. The sum over the units from hidden layer 2 can hence be performed analytically.

2. Perform MC-sampling for the hidden nodes from hidden layer 1.

This