Title: Date:

11. Artificial intelligence and quantum physics

The learning goals for the last two weeks are:

* You know what an artificial neural net is.
* You know the difference between supervised and unsupervised learning.
* You can give examples of how supervised and unsupervised learning can be used to determine phase transitions.
* You know what a restricted Boltzmann machine is.
* You can elaborate how such an RBM can be used to represent wave functions

Title: Date:

11.2 Quantum states from neural nets

Today's lecture is based on the following papers

1. Carleo & Troyer Science <u>355</u>, 602 (2017)

2. Kaubrügger et al. Phys. Rev. B <u>97</u> 195136 (2018)

3. Torlai et al. Nature Phys. <u>14</u> 447 (2018)

We are going to learn how an artificial neural net can be used to represent a quantum wave-function.

11.2.1 Restricted Boltzmann machines

The probability distribution $q(\{s_i\})$ of N variables s_i is unknown but shall be learned from samples drawn from the unknown $q(\{s_i\})$.

We assume that we can write

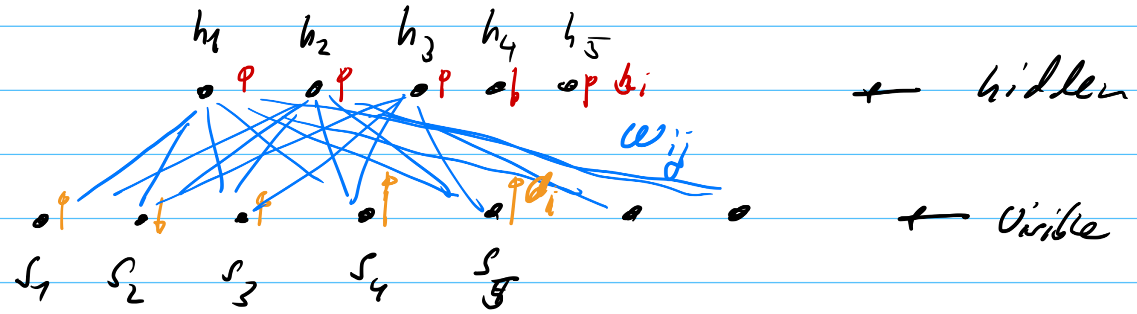
$$q(\{s_i\}) \approx p(\{s_i\}) = e^{-E(\{s_i\})}$$

and hence we call it a Boltzmann machine.

Title: Date:

Specifically, we write

$$E(\{s_i\}) = \sum_i s_i a_i + \sum_{ij} w_{ij} s_i h_j + \sum_i b_i h_i$$



Our task is to learn $\{w_{ij}, b_i, a_i\}$ such that

$$p(\{s_i\}) = \sum_{h_i = \pm} e^{-E(\{s_i\})}$$

represents the samples drawn from $q(\{s_i\})$ as well as possible.

A breakthrough came with the use of contrastive divergence by Hinton, which enables an efficient training of the weights. "On contrastive divergence learning" by Carreira-Perpiñán & Hinton.

For quantum wave functions we can do something else. We don't need to rely on samples (i.e. configurations sampled via Monte-Carlo) but we can capitalize on the fact that the wave-function itself has a simple cost-function

$$E_{var} = \frac{\langle \psi_{\vec{\omega}} | H | \psi_{\vec{\omega}} \rangle}{\langle \psi_{\vec{\omega}} | \psi_{\vec{\omega}} \rangle}$$

This allows us to do an imaginary time evolution

$$|\psi_{\omega}^{\tau}\rangle = e^{-\tau H} |\psi_{\omega}^{0}\rangle \quad \text{with}$$

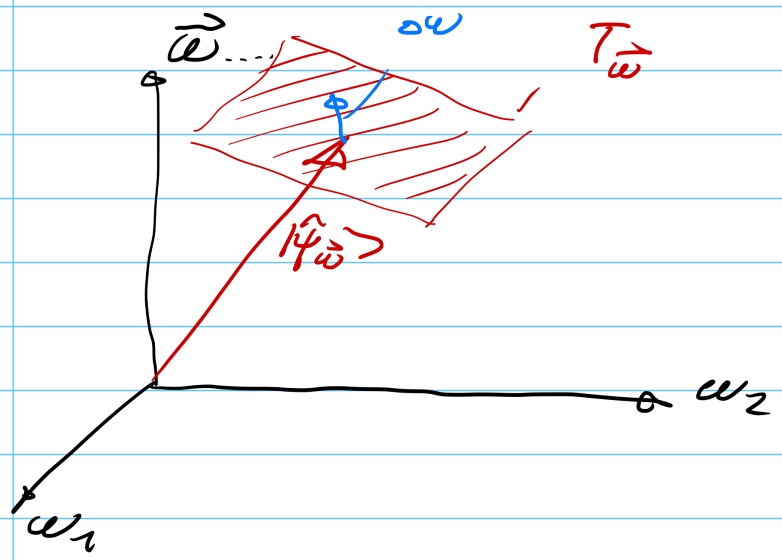
$\tau \to \infty$ to find the optimal $\{\omega_{ij}, h_i, a_i\}$.

11.2.2. Stochastic reconfiguration

We have the normalized wave function

$$|\hat{\psi}_{\vec{\omega}}\rangle,$$

where $\vec{\omega}$ contains all parameters $\{\omega_{ij}, h_i, a_i\}$.



In an update step, we would like to move in the tangent space $T_{\vec{\omega}}$ to our current wave function. T_{ω} is spanned by

Title: Date:

$$|f_{\vec{\omega}}\rangle = |\partial_{\omega_j}\hat{\psi}_{\vec{\omega}}\rangle - |\hat{\psi}_{\omega}\rangle\langle\hat{\psi}_{\omega}|\partial_{\omega_j}|\hat{\psi}_{\vec{\omega}}\rangle$$

Note, that the $|f_{\omega_j}\rangle$ are not orthogonal.

We now write

$$|\hat{\psi}_{\vec{\omega}}(\tau+d\tau)\rangle = e^{-d\tau H}\,|\hat{\psi}_{\vec{\omega}}(\tau)\rangle$$

and expand the left-hand side

$$|\psi_{\vec{\omega}}(\tau+d\tau)\rangle \approx |\hat{\psi}_{\vec{\omega}}(\tau)\rangle + d\tau\sum_{k=1}^{N_{\vec{\omega}}}\dot{\omega}_k(\tau)\times$$
$$\left[|\partial_{\omega_k}\hat{\psi}_{\vec{\omega}}(\tau)\rangle - |\hat{\psi}_{\vec{\omega}}(\tau)\rangle\langle\hat{\psi}_{\vec{\omega}}(\tau)|\right.$$
$$\left.|\partial_{\omega_k}\hat{\psi}_{\omega}(\tau)\rangle\right]$$

where we subtracted the last term to ensure we move perpendicular to $|\hat{\psi}_{\vec{\omega}}(\tau)\rangle$.

Expanding also the right-hand side and multiplying from the left with $\langle f_{\omega}|$ we obtain

$$\sum_{k=1}^{N_{\vec{\omega}}}\dot{\omega}_k\left[\langle\partial_{\omega_j}\hat{\psi}_{\vec{\omega}}|\partial_{\omega_k}\hat{\psi}_{\omega}\rangle - \langle\partial_{\omega_j}\hat{\psi}_{\vec{\omega}}|\hat{\psi}_{\vec{\omega}}\rangle\langle\hat{\psi}_{\vec{\omega}}|\partial_{\omega_k}\hat{\psi}_{\omega}\rangle\right]$$

$$= -\langle\partial_{\omega_j}\hat{\psi}_{\vec{\omega}}|H|\hat{\psi}_{\vec{\omega}}\rangle + \langle\partial_{\omega_j}\hat{\psi}_{\omega}|\hat{\psi}_{\vec{\omega}}\rangle\langle\hat{\psi}_{\omega}|H|\hat{\psi}_{\vec{\omega}}\rangle$$

Title: Date:

or

$$S_{\vec{w}} \frac{d\vec{w}}{d\tau} = - \vec{F}_{\vec{w}}$$

with

$$S_{\vec{w}} = \langle \partial_{w_j} \hat{\psi}_{\vec{w}} | \partial_{w_i} \hat{\psi}_{\vec{w}} \rangle - \langle \partial_{w_j} \hat{\psi}_{\vec{w}} | \hat{\psi}_{\vec{w}} \rangle \langle \hat{\psi}_{\vec{w}} | \partial_{w_i} \hat{\psi}_{\vec{w}} \rangle$$

the metric or co-variance tensor and

$$\vec{F}_{\vec{w}} = \langle \partial_{w_j} \hat{\psi}_{\vec{w}} | H | \hat{\psi}_{\vec{w}} \rangle - \langle \partial_{w_j} \hat{\psi}_{\vec{w}} | \hat{\psi}_{\vec{w}} \rangle \langle \hat{\psi}_{\vec{w}} | H | \hat{\psi}_{\vec{w}} \rangle$$

by introducing γ as the imaginary time step we find

$$\boxed{d\vec{w} = -\gamma\, S_{w}^{-1} \vec{F}_{\vec{w}}}$$

We now need to estimate $\vec{F}_{\vec{w}}$ and $\vec{S}_{w}$ from Monte-Carlo sampling from

$$\langle \hat{\psi}_{\vec{w}} | \hat{\psi}_{\vec{w}} \rangle .$$

See Kaubrügger et al. Phys. Rev. B 97 195136 (2018) and references therein for the details how to do so.