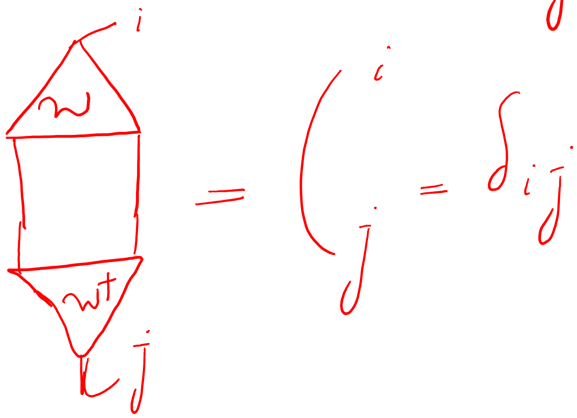
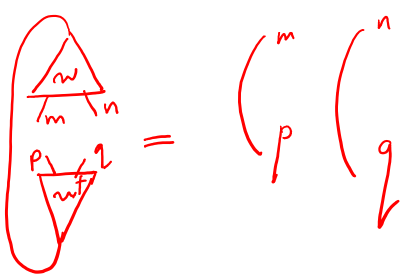
Preliminary: |i| shows bond dimension of index "i". Tensors "w" and "u" are unitary and are represented as follows:

(1) $w^{\dagger} w = I \Rightarrow (w^{\dagger} w)_{ij} = \delta_{ij}$

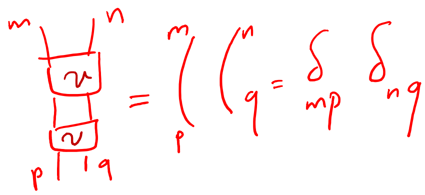


$= \delta_{ij}$

(2) $w w^{\dagger} = I \Rightarrow (w w^{\dagger})_{mn;pq} = \delta_{mp} \delta_{nq}$

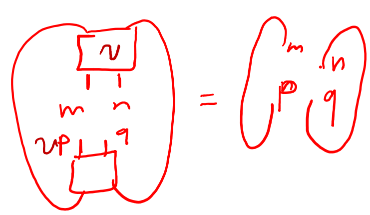


(3) $u^{\dagger} u = I \Rightarrow$



$= \delta_{mp} \delta_{nq}$

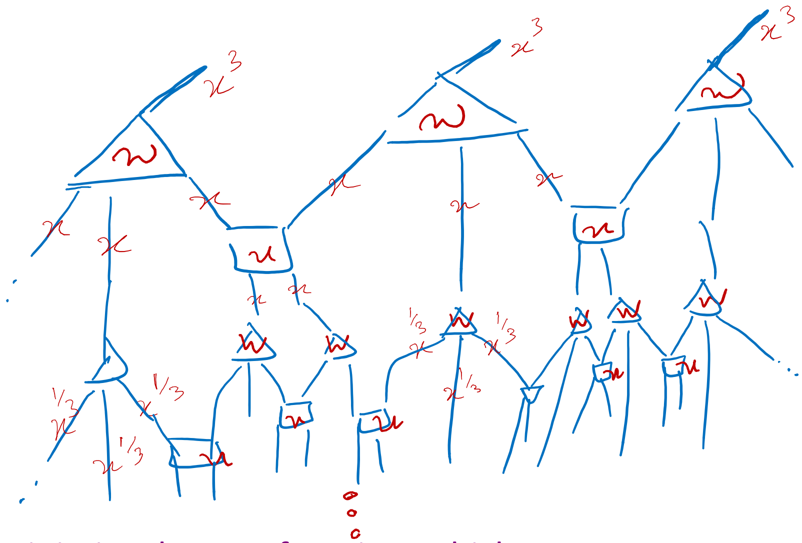
$u u^{\dagger} = I \Rightarrow$



$= \delta_{mp} \delta_{nq}$

Let assume the unitary "$\hat{\mathcal{U}}$", which is supposed to diagonalize the many-body localized Hamiltonian, possesses a unitary ternary MERA structure, where bond dimension at last layer is equal to χ^3.

$$\hat{\mathcal{U}}_{mera} =$$



Main goal is to minimize the cost function, which was introduced by you, but here instead we use MERA structure and also its optimization techniques, let me focus on optimizing the cost function:

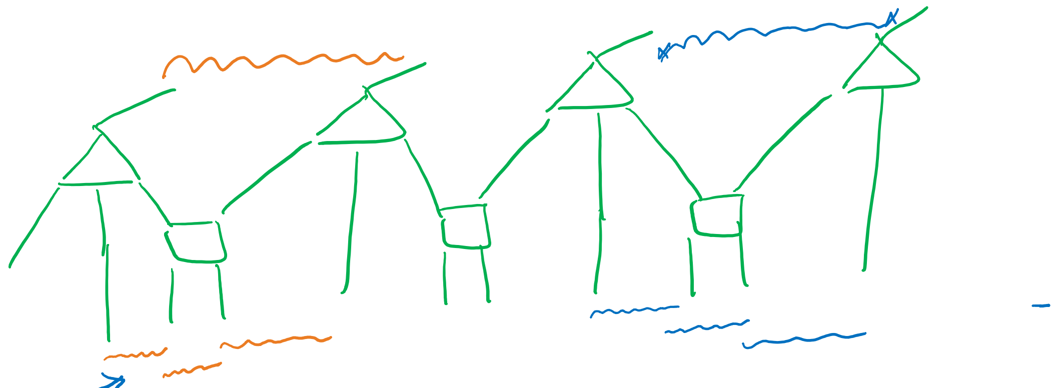
$$f(\{w\},\{u\}) = \sum_{\{J\}} \langle\psi_J|H^2|\psi_J\rangle - \langle\psi_J|H|\psi_J\rangle^2 \geq 0$$

$$= tr(H^2) - \sum_{\{J\}} \langle\psi_J|H|\psi_J\rangle^2 \geq 0$$

Only second term is dependent on unitary tensors "u" and "w", so one just needs to maximize the second term.

$$\max_{\{w\}\{u\}} \sum_{\{J\}} \langle\psi_J|H|\psi_J\rangle^2 \geq 0$$

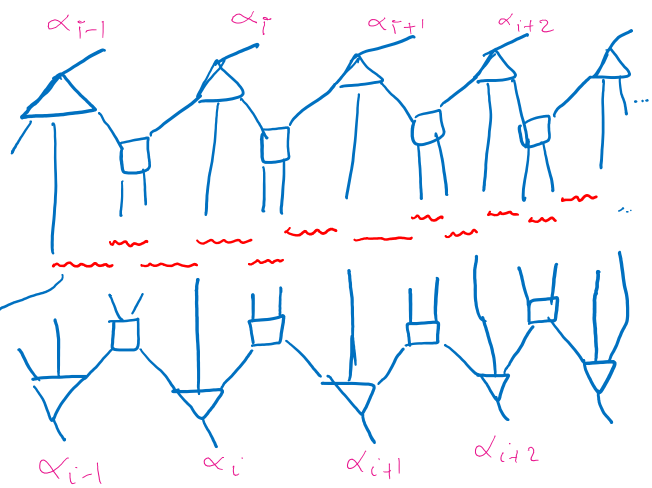
Goal here is to take advantage of locality of Hamiltonian and simplify the cost function as simple as it's possible. Point here is that MERA preserves locality of Hamiltonian, i.e. at each layer, the interactions remain local, e.g. two-body interaction at the first layer is being mapped into two-body interaction on the super spin at the second layer. It helps us to simplify the cost function.



two-body interactions are being mapped into two body interaction on super spin at higher layer.

Let just focus on optimizing the tensors that are at the last layer, and

$$\langle \psi_J | H | \psi_J \rangle = \cdots$$



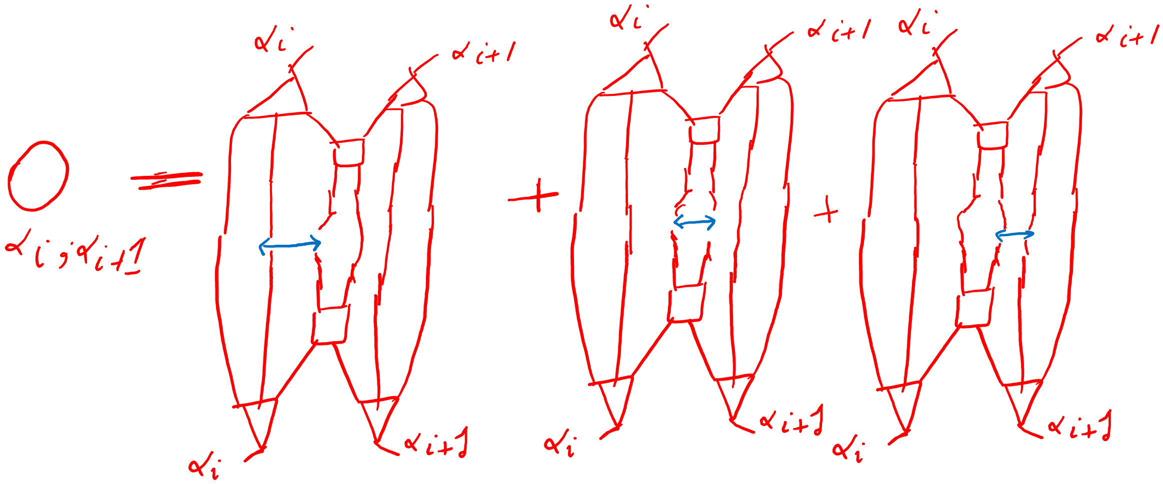
two-body interactions

$\sim\sim\sim = \otimes$

Note: $\{J\} = \{\alpha_1, \ldots, \alpha_N\} \Rightarrow |\psi_{\{J\}}\rangle = |\psi\rangle_{\{\alpha_1 \ldots \alpha_N\}}$

By using properties of unitary, one comes to the following result:

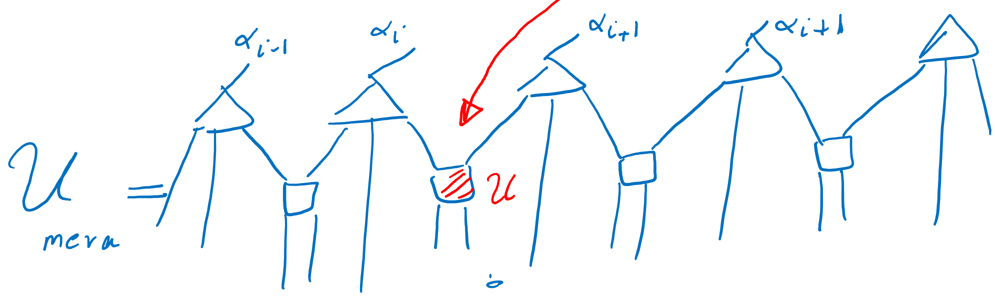
$$\langle \psi_J | H | \psi_J \rangle = \sum_{i=1}^{N-1} O_{\alpha_i ; \alpha_{i+1}}$$



So we have:

$$\sum_J \langle \psi_J | H | \psi_J \rangle^2 = \sum_{\alpha_1 \cdots \alpha_N} \left(\sum_{i=1}^{N-1} O_{\alpha_i;\alpha_{i+1}} \right)^2 \qquad (1)$$

Now, my goal is to obtain e.g. the unitary "u", as depicted in the following figure, to maximize the cost function; I aim to describe the method:



$$\sum_J \langle \psi_J | H | \psi_J \rangle^2 = \sum_{\alpha_1 \cdots \alpha_N} \left(\cdots + O_{\alpha_i;\alpha_{i+1}}(u) + \cdots \right) \left(\cdots + O_{\alpha_i;\alpha_{i+1}}(u) + \cdots \right)$$

Notice that only $O_{\alpha_i;\alpha_{i+1}}$ is dependent on $\mathcal{U}$, all other

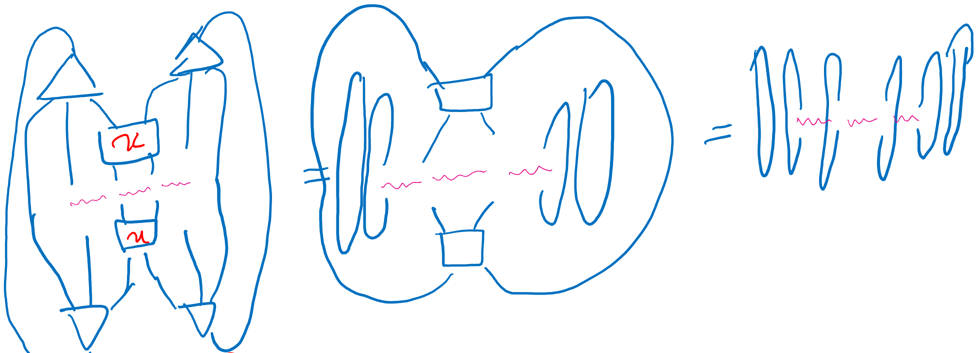
terms are independent of it.

$$= \sum_{\alpha_1 \cdots \alpha_N} \cdots + 2 O_{\alpha_{i-1};\alpha_i} O_{\alpha_i;\alpha_{i+1}}(u) + O^2_{\alpha_i;\alpha_{i+1}}(u) + 2 O_{\alpha_i;\alpha_{i+1}}(u)\, O_{\alpha_{i+1};\alpha_{i+2}} + \cdots$$

Only above three terms are dependent on $\mathcal{U}$, and all other terms becomes independent of it!!!!; the reason comes from the following: e.g. consider the following term:

$$\sum_{\alpha_1\ldots\alpha_N} O^{(u)}_{\alpha_i,\alpha_{i+1}} O_{\alpha_{i+2};\alpha_{i+3}} = |\alpha|^{N-4} \times \Big(\sum_{\substack{\alpha_i\\ \alpha_{i+1}}} O^{(u)}_{\alpha_i;\alpha_{i+1}}\Big)\Big(\sum_{\substack{\alpha_{i+2}\\ \alpha_{i+3}}} O_{\alpha_{i+2};\alpha_{i+3}}\Big).$$

$$\sum_{\alpha_i\alpha_{i+1}} O^{(u)}_{\alpha_i,\alpha_{i+1}} =$$



Which is totally independent of u. Therefore, only the following three terms play role in cost function, regarding optimization of unitary matrix u. So, one just needs to consider following tensor network contraction and use MERA techniques to linearize it and then obtain desired result.

$$\sum_{\substack{\alpha_i,\alpha_{i+1}\\ \alpha_{i+2},\alpha_{i-1}}} \Big(2 O_{\alpha_{i-1},\alpha_i} O^{(u)}_{\alpha_i,\alpha_{i+1}} + O^{2(u)}_{\alpha_i,\alpha_{i+1}} + 2 O^{(u)}_{\alpha_i,\alpha_{i+1}} O_{\alpha_{i+1},\alpha_{i+2}}\Big)$$

The above expression could be written as coming,

$$= tr(Y_u u)$$

Where Y_u is environment of unitary matrix u.

To optimize it, we use singular value decomposition and express Y_u :

$$Y_u = M S N^\dagger \;;\; M^\dagger M = I \text{ and } N^\dagger N = I$$

Now, we choose $\mathcal{U}$:

$$\mathcal{U} = N M^{\dagger} \qquad (2)$$

$$\Longrightarrow tr(\mathcal{Y}_{\mathcal{U}} \mathcal{U}) = (M S N^{\dagger} N M^{\dagger}) = tr(S) \geq 0$$

Now, after this step, we again calculate environment $\mathcal{Y}_{\mathcal{U}}$ and iteratively continue this process to converge to a specific value.

Let me again come back to calculation of environment and show the explicit form of its tensor network contraction.

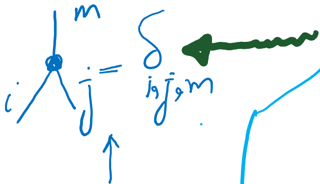
$$\text{Cost function} = \sum_{\substack{\alpha_{i-1} \\ \alpha_i \\ \alpha_{i+1} \\ \alpha_{i+2}}} \left(O_{\alpha_{i-1},\alpha_i} + O(u)_{\alpha_i,\alpha_{i+1}} + O_{\alpha_{i+1},\alpha_{i+2}} \right) O(u)_{\alpha_i,\alpha_{i+1}}$$

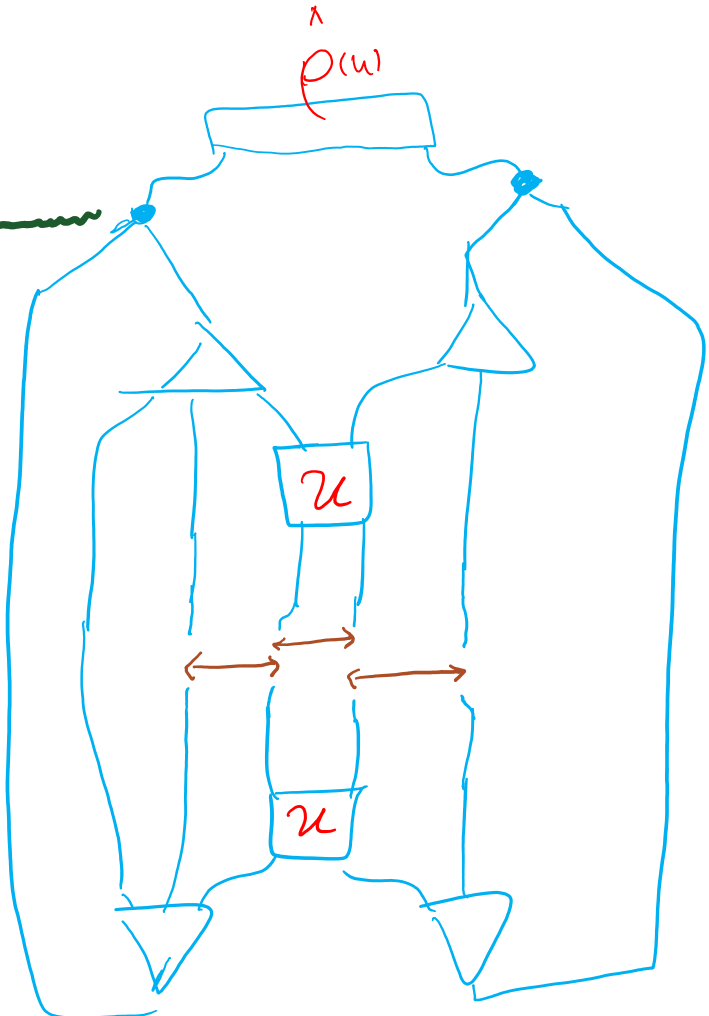
$$= \sum_{\alpha_i,\alpha_{i+1}} \underbrace{\hat{\rho}(u)_{\alpha_i,\alpha_{i+1}}} \, O(u)_{\alpha_i,\alpha_{i+1}}$$

$$\hat{\rho}(u)_{\alpha_i,\alpha_{i+1}} = \sum_{\substack{\alpha_{i-1} \\ \alpha_{i+2}}} \left(O_{\alpha_{i-1},\alpha_i} + O(u)_{\alpha_i,\alpha_{i+1}} + O_{\alpha_{i+1},\alpha_{i+2}} \right)$$

$$= |\alpha| \sum_{\alpha_{i-1}} O_{\alpha_{i-1},\alpha_i} + |\alpha|^2 \, O(u)_{\alpha_i,\alpha_{i+1}} + |\alpha| \sum_{\alpha_{i+2}} O_{\alpha_{i+1},\alpha_{i+2}}$$

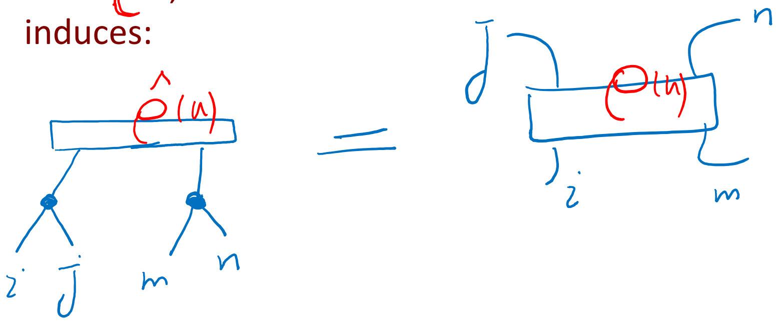
Now, the cost function gets a simple form, which is similar to that of MERA contraction for getting ground state:



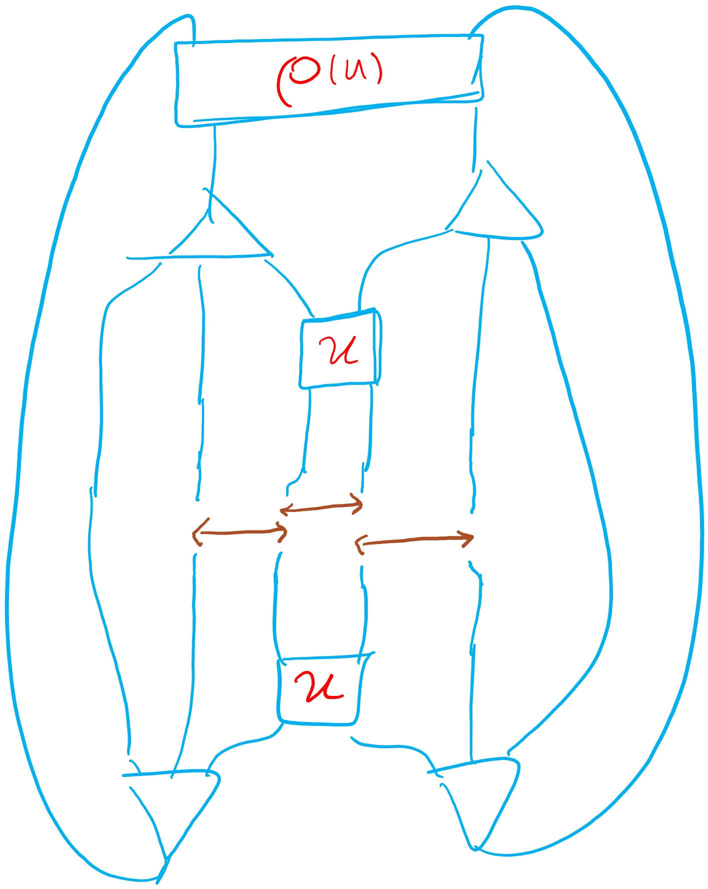


It means these three induces are equal to each other, otherwise It is zero

Now let show the following tensor with ρ , where we have absorbed some induces:

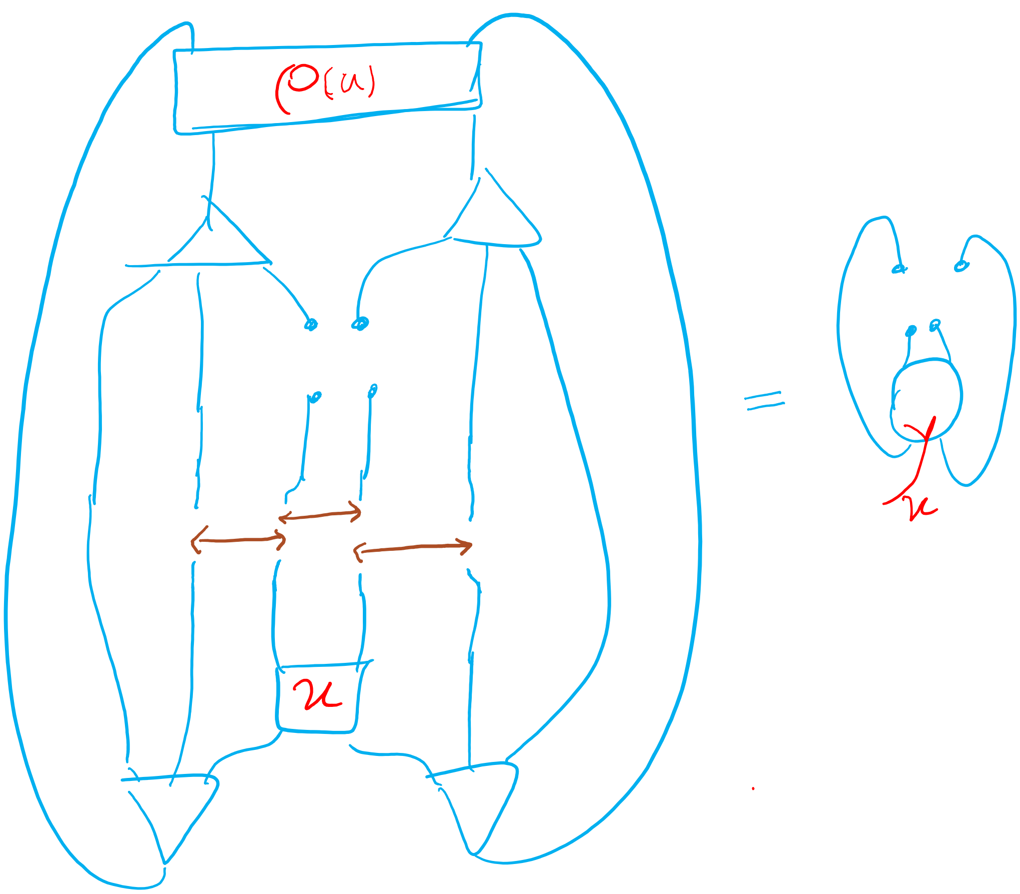


So, we finally come to the following result:

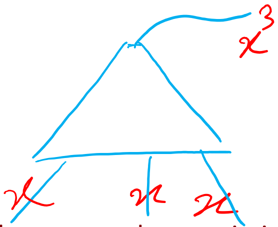


(A^*)

For calculation of environment, one just needs to use the following tensor network contraction:



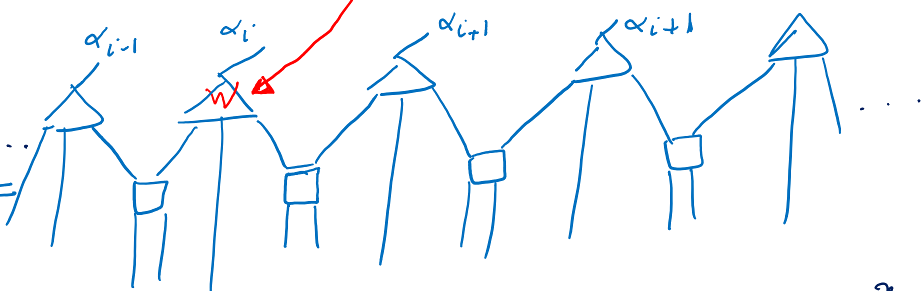
Complexity of contraction scales as χ^{12}, where χ is bond dimension of tensors as comes:



Other tensors could be managed to maximize the cost function in a similar way, the point here is that the tensors at low layers even could be easily optimized by using MERA techniques, i.e. the so called Descending operator, which I'll explain it.

Now, let obtain unitary "w", as depicted in the following figure, to maximize the cost function:

$\mathcal{U}_{mera} =$



$$\sum_{j} \langle\psi_j|H|\psi_j\rangle^2 = \sum_{\alpha_1\ldots\alpha_N} \left(\cdots + O_{\alpha_{i-2},\alpha_{i-1}} + O_{\alpha_{i-1},\alpha_i}(w) + O_{\alpha_i,\alpha_{i+1}}(w) + O_{\alpha_{i+1},\alpha_{i+2}} + \cdots\right)^2$$

$$= \sum_{\alpha_1\ldots\alpha_N} \cdots + O^2_{\alpha_{i-1},\alpha_i}(w) + 2\, O_{\alpha_{i-1},\alpha_i}(w)\, O_{\alpha_i,\alpha_{i+1}}(w) + O^2_{\alpha_i,\alpha_{i+1}}(w) + \cdots$$

$$= \sum_{\substack{\alpha_{i-1}\\ \alpha_i\\ \alpha_{i+1}}} \left(O^2_{\alpha_{i-1},\alpha_i}(w) + O^2_{\alpha_i,\alpha_{i+1}}(w) + 2\, O_{\alpha_{i-1},\alpha_i}(w)\, O_{\alpha_i,\alpha_{i+1}}(w) \right) + \text{other terms}$$

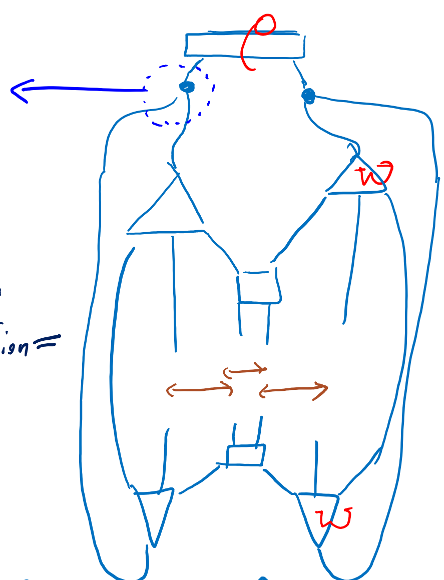
These terms are independent of unitary "w"

$$\Rightarrow \text{Cost function} = \sum_{\substack{\alpha_i \\ \alpha_{i+1}}} \rho_{\alpha_i,\alpha_{i+1}} \, O^{(w)}_{\alpha_i,\alpha_{i+1}}$$

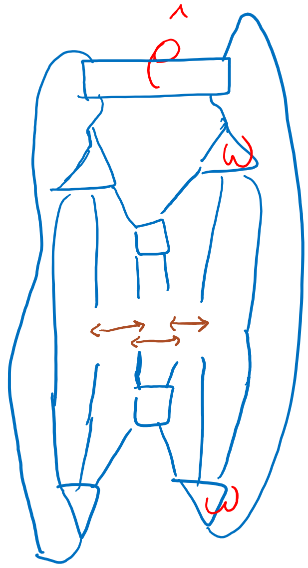
$$\rho_{\alpha_i,\alpha_{i+1}} = \sum_{\alpha_{i-1}} \left(O^{2\,(w)}_{\alpha_{i-1},\alpha_i} + O_{\alpha_i,\alpha_{i+1}} + 2\, O_{\alpha_{i-1},\alpha_i} \right)$$

delta function

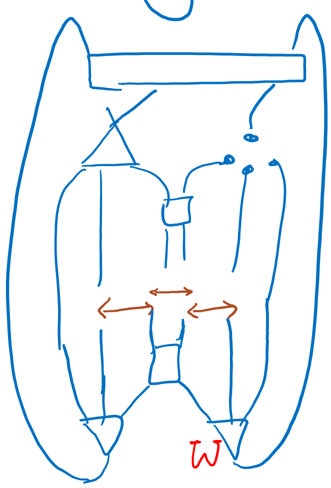
$\Rightarrow$ Cost function =



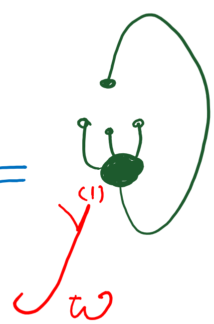
=



$\Rightarrow$



=



= Environment of w

ENVIRONMENT OF "W"

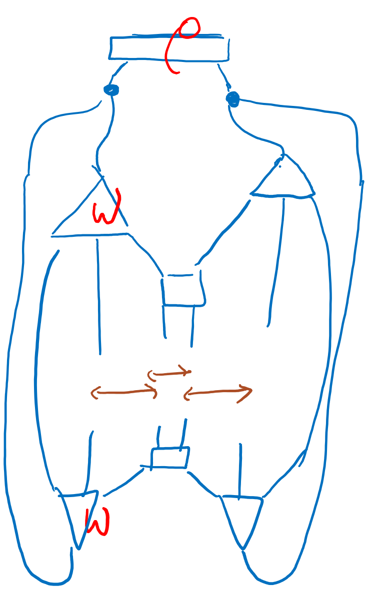
One could also find another expression for environment of "w", which makes the environment more reliable and It seems it's better to consider both of them, let me show graphically what I mean:

$$\text{Cost function} = \sum_{\substack{\alpha_{i-1} \\ \alpha_i \\ \alpha_{i+1}}} \left(O^{2\,(w)}_{\alpha_{i-1},\alpha_{i-1}} + O^{2\,(w)}_{\alpha_i,\alpha_{i+1}} + 2\, O^{(w)}_{\alpha_{i-1},\alpha_i} \, O^{(w)}_{\alpha_i,\alpha_{i+1}} \right)$$

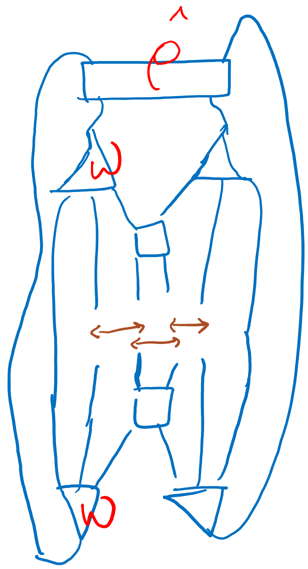
$$= \sum \rho^{(\omega)}_{\alpha_{i-1},\alpha_i} \, O_{\alpha_{i-1},\alpha_i} \quad .$$

$$\rho^{(\omega)}_{\alpha_{i-1},\alpha_i} = \sum_{\alpha_{i+1}} \left(O^{(\omega)}_{\alpha_{i-1},\alpha_i} + \overset{2}{O}{}^{(\omega)}_{\alpha_i,\alpha_{i+1}} + z\, O_{\alpha_i,\alpha_{i+1}} \right)$$

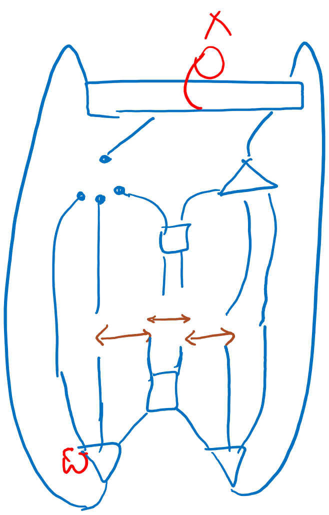
Cost function =



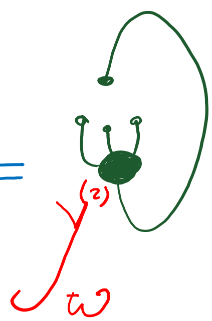
=



⇒



=



= Environment of ω

ENVIRONMENT OF "W"

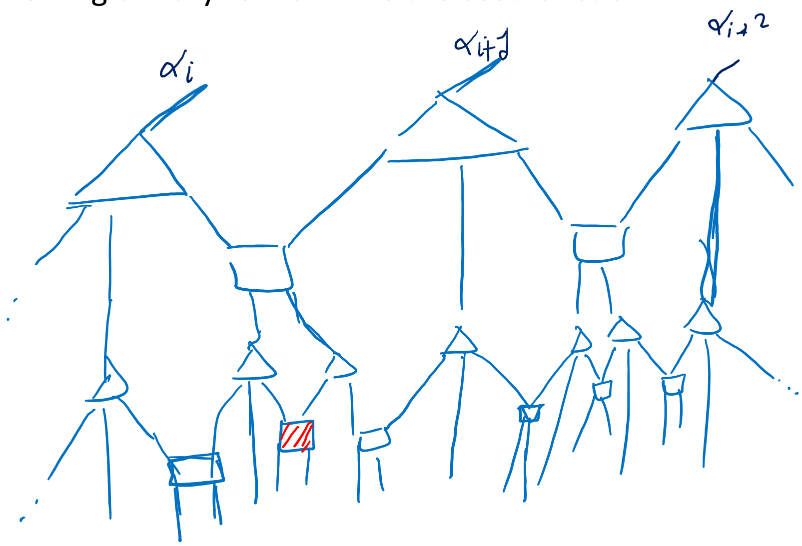
Note I've obtained two environments i.e. $Y_w^{(1)}, Y_w^{(2)}$; I sum up both of them to obtain unitary "w": (1) doing so, makes environment more symmetric and seems that it results in more accurate result---in case of MERA for ground state. see the following:

$$\text{Cost function} = tr\left(\left(Y_w^{(1)} + Y_w^{(2)}\right) w\right) / 2$$

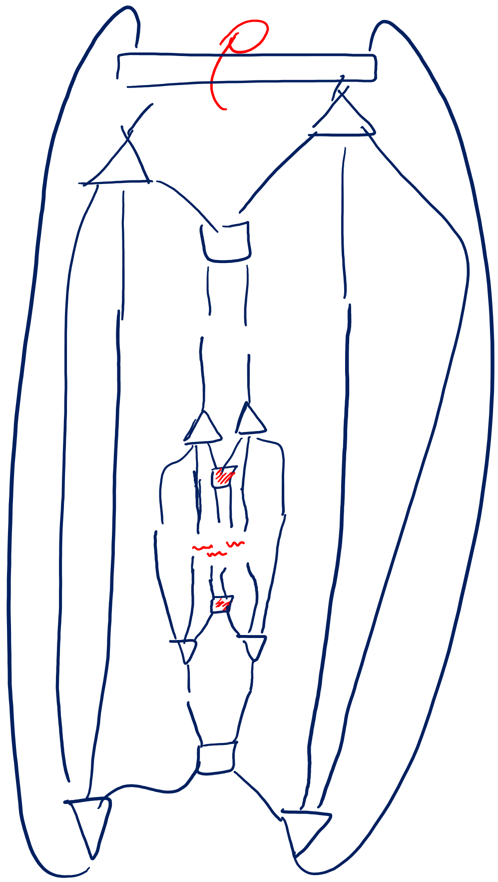
$$\Rightarrow \frac{Y_w^{(1)} + Y_w^{(2)}}{2} = M S N^{\dagger} \Rightarrow \boxed{w = N M}$$

Obtaining unitary tensors at lower layers are also straightforward, and could be done by relying on Descending operator, which are similar to that of MERA for ground state, let me explain it in more detail, suppose I want to obtain the following unitary to maximize the cost function:

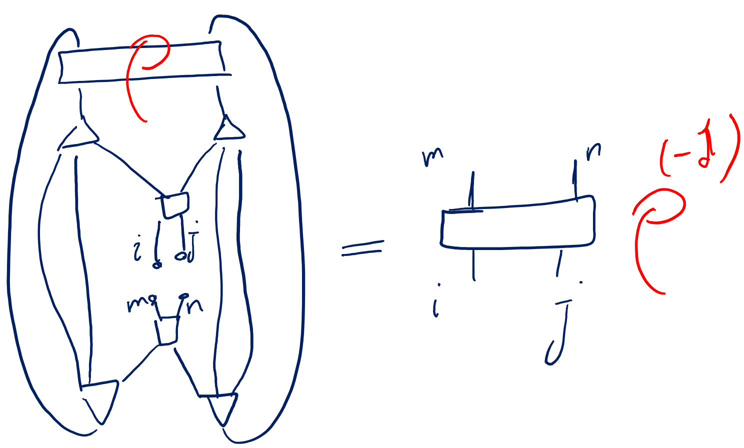
$\hat{u}_{mera} =$



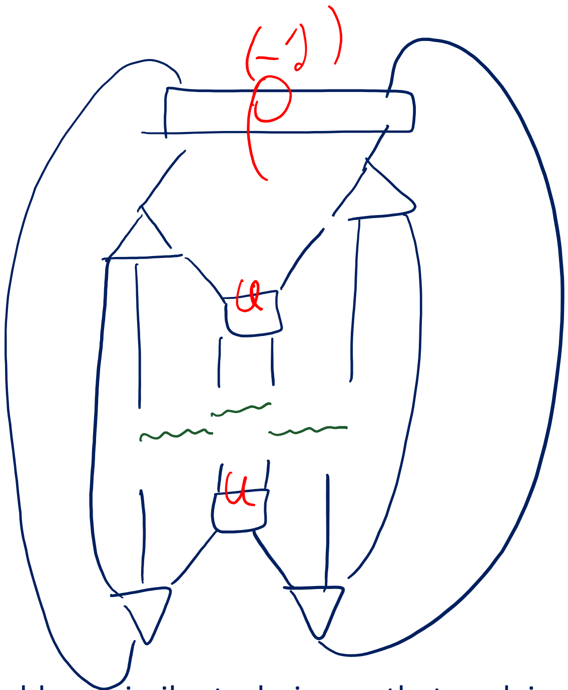
cost function =



Please note I've used the equation  ,
now, I use "descending operator" to get density matrix at lower layers:



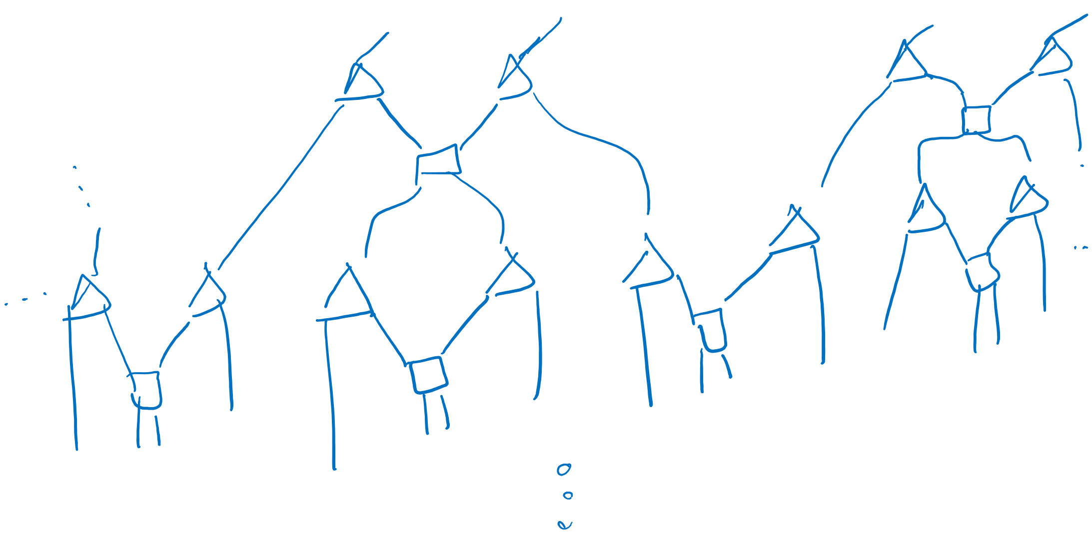
Now, I could easily obtain environment of "u":



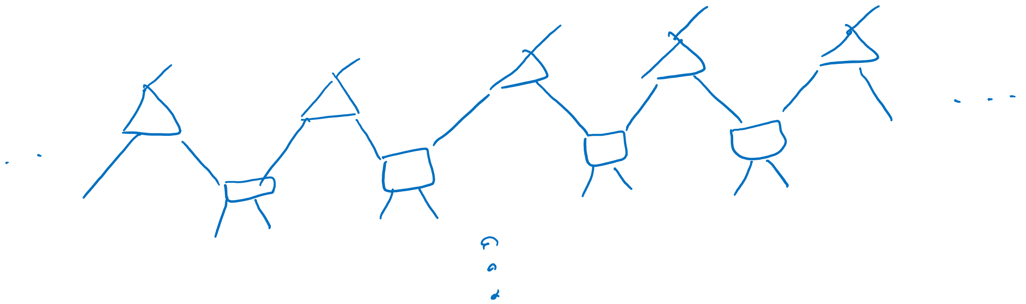
Here, we could use similar techniques, that explained before in this note, to get 'u'. The same also could be done for other tensors: one just needs to rely on "Descending operators". Note optimizing this part of tensors is not considerable---in point view of computational resources---in comparison with that of the last layer, since we are dealing with bond dimension $\chi^{1/3}$.

Other schemes of MERA is also possible that have different complexity and entanglement (between two arbitrary cut), for example see the below figures and table:

Modified binary MERA:

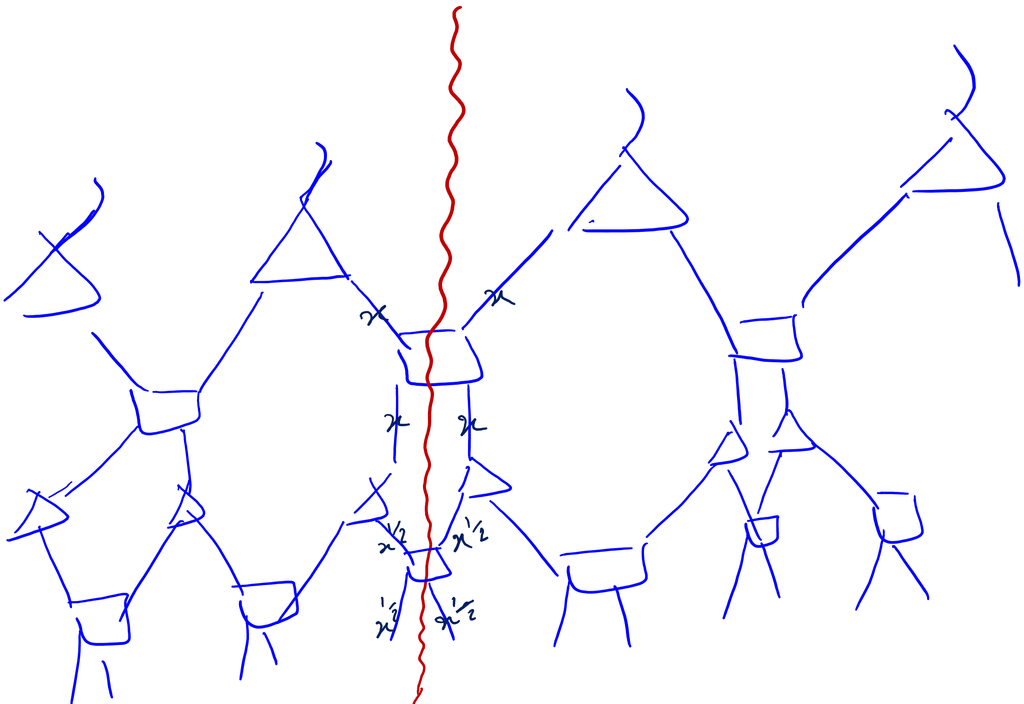


Binary MERA:



method	Complexity	Entanglement
Modified Binary MERA	$O(L\chi^{9})$	$\log(\chi^{3/2})$
Ternary MERA	$O(L\chi^{12})$	$\log(\chi^{2})$
Binary MERA	$O(L\chi^{12})$	$\log(\chi^{2})$

By Entanglement, I suppose e.g. the following cut:



Entanglement $\sim \log x^3$