What we want to prove is that

$$\sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W(q,\hat{q}) \leq \epsilon_3 \tag{1}$$

Note this is slightly different from the lemma in the original paper, but if we look at where it is used this is all we need. An easy consequence would be that for all q
$\sum_{\hat{q}:\Phi(q)\neq\hat{q}} W(q,\hat{q}) \leq \epsilon_3$.

0.1 Tail

First we show that the amount of weight greater than a certain distance can be bounded, which is kind of trivial. We know that by the bound on the expected length.

$$\sum_{d\geq 0} W_d \leq L+1 \tag{2}$$

or

$$\sum_{d>0} W_d \leq L \tag{3}$$

so since W_k decreases monotonically for all $k>0$

$$\sum_{d>0}^{k} W_k \leq L \tag{4}$$

i.e.

$$W_k \leq \frac{L}{k} \tag{5}$$

From this we can conclude that

$$\sum_{d>k} W_k \leq \frac{L^2}{k} \tag{6}$$

0.2 Lemma

We just want to show that the total weight not in the right states is less than dx for some constant $x = |\Sigma|\epsilon_5 + n\epsilon_2$.

i.e. we want to prove that for all $d \geq 0$

$$\sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W_d(q,\hat{q}) \leq dx \tag{7}$$

We will do this by induction on d. Note that for $d=0$, $W_0(q_0,\Phi(q_0)) = W_0(q_0) = 1$. Assume that it is true for $d-1$.

Recall that by definition of $W_d(q,\hat{q})$.

$$W_d(q,\hat{q}) = \sum_{p\in Q}\sum_{\hat{p}\in\hat{Q}} W_{d-1}(p,\hat{p}) \sum_{\substack{\sigma:\tau(p,\sigma)=q\\ \hat{\tau}(\hat{p},\sigma)=\hat{q}}} \gamma(p,\sigma)$$

Substituting this in we get the following quadruple sum.

$$\sum_{q}\sum_{\hat{q}:\Phi(q)\neq\hat{q}}\sum_{p\in Q}\sum_{\hat{p}\in\hat{Q}} W_{d-1}(p,\hat{p}) \sum_{\substack{\sigma:\tau(p,\sigma)=q\\ \hat{\tau}(\hat{p},\sigma)=\hat{q}}} \gamma(p,\sigma)$$

Looking at the quadruple sum we can divide them into three bits. First where $\Phi(p)\neq\hat{p}$. These correspond to places where the weight is in the wrong state at $d-1$.

$$\sum_{q}\sum_{\hat{q}:\Phi(q)\neq\hat{q}}\sum_{p\in Q}\sum_{\hat{p}\neq\Phi(p)} W_{d-1}(p,\hat{p}) \sum_{\substack{\sigma:\tau(p,\sigma)=q\\ \hat{\tau}(\hat{p},\sigma)=\hat{q}}} \gamma(p,\sigma)$$

Interchanging the order of summation this will be

$$\sum_{p\in Q}\sum_{\hat{p}\neq\Phi(p)} W_{d-1}(p,\hat{p}) \sum_{q}\sum_{\hat{q}:\Phi(q)\neq\hat{q}} \sum_{\substack{\sigma:\tau(p,\sigma)=q\\ \hat{\tau}(\hat{p},\sigma)=\hat{q}}} \gamma(p,\sigma)$$

Using the fact that for each $\sigma, p, \hat{p}$ there is a unique pair of states such that $\tau(p,\sigma)=q, \hat{\tau}(\hat{p},\sigma)=\hat{q}$ we can see that this will be less than (using the inductive assumption)

$$\sum_{p\in Q}\sum_{\hat{p}\neq\Phi(p)} W_{d-1}(p,\hat{p}) \sum_{\sigma}\gamma(p,\sigma) \leq (d-1)x$$

The remaining part will be where $\hat{p}=\Phi(p)$. (i.e the weight was in the right state at $d-1$). We can consider first cases where $W(p)<\epsilon_2$. The sum will then be bounded by $n\epsilon_2$.

The final case is where $W(p)\geq\epsilon_2$. In this case we know that $\gamma(p,\sigma)<\epsilon_5$. So the sum is less than

$$\sum_{q}\sum_{\hat{q}:\Phi(q)\neq\hat{q}}\sum_{p\in Q} W_{d-1}(p,\hat{p}) \sum_{\substack{\sigma:\tau(p,\sigma)=q\\ \hat{\tau}(\hat{p},\sigma)=\hat{q}}} \epsilon_5$$

Interchanging again

$$\sum_{p\in Q} W_{d-1}(p,\hat{p}) \sum_{\sigma}\epsilon_5 \leq |\Sigma|\epsilon_5$$

This establishes that

$$\sum_{q}\sum_{\hat{q}:\Phi(q)\neq\hat{q}} W_d(q,\hat{q}) \leq (d-1)x + |\Sigma|\epsilon_5 + n\epsilon_2 \tag{8}$$

QED.

0.3 Conclusion

Now we set

$$D = \frac{2L^2}{\epsilon_3} \tag{9}$$

We can now show that

$$\sum_d \sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W_d(q,\hat{q}) \leq \sum_{d\leq D} \sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W_d(q,\hat{q}) + \sum_{d>D} W_d \tag{10}$$

$$\sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W(q,\hat{q}) \leq \sum_{d\leq D} dx + LW_{D-1} \tag{11}$$

$$\sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W(q,\hat{q}) \leq \frac{1}{2}xD(D+1) + \frac{L^2}{D} \leq xD^2 + \frac{L^2}{D} \tag{12}$$

So now we substitute in x and D

$$\sum_q \sum_{\hat{q}:\Phi(q)\neq\hat{q}} W(q,\hat{q}) \leq (|\Sigma|\epsilon_5 + n\epsilon_2)\frac{4L^4}{\epsilon_3^2} + \frac{\epsilon_3}{2} \tag{13}$$

We now set

$$\epsilon_2 = \frac{\epsilon_3^3}{8nL^4} \tag{14}$$

and

$$\epsilon_5 = \frac{\epsilon_3^3}{8|\Sigma|L^4} \tag{15}$$

This establishes the result. Note that these quantities ϵ_2, ϵ_5 are much tighter bound than before. This will have an increase in the sample complexity.

0.4 Consequences

This needs some changes to the rest of the paper. The algorithm is fine, and the result is fine, but some of the bounds need to be modified.

In Section 5 when we bound D_2 it will have

$$D_2 \leq \epsilon_3 \log \frac{1}{\gamma_{min}} \tag{16}$$

So the definition of ϵ_3 can become (a bit larger):

$$\epsilon_3 = \frac{\epsilon}{4\log(4(L+1)(|\Sigma|+1)/\epsilon)} \tag{17}$$

$$\epsilon_6 = \frac{\epsilon_2\epsilon_5}{(L+1)} \tag{18}$$

N will need to be changed to

$$N = \frac{2n|\Sigma|m_0}{\epsilon_6} \tag{19}$$