Approximation Accuracy of MPS and Renyi Entropy

Ken Shiozaki

June 1, 2026

Abstract

This is a calculation memo for [1].

Consider an N-site system. Let the dimension of the local Hilbert space be d. Take the MPS in canonical form:

$$\left|\psi\right\rangle = \sum_{i_1,\dots,i_N} A^{[1]}_{i_1} A^{[2]}_{i_2} \cdots A^{[N-1]}_{i_{N-1}} A^{[N]}_{i_N} \left|i_1 \cdots i_N\right\rangle. \tag{1}$$

Here $A^{[m]}_i$ is a $D_m \times D_{m+1}$ matrix and $D_1 = D_{N+1} = 1$. They satisfy

$$\sum_i A^{[m]}_i A^{[m]\dagger}_i = 1_{D_m}, \quad 1 \leq m \leq N, \tag{2}$$

$$\sum_i A^{[m]\dagger}_i \Lambda^{[m-1]} A^{[m]}_i = \Lambda^{[m]}, \quad 1 \leq m \leq N, \tag{3}$$

$$\Lambda^{[0]} = \Lambda^{[N]} = 1, \quad \Lambda^{[m]} = \mathrm{diag}(\lambda^{[m]}_1, \dots, \lambda^{[m]}_{D_m}), \quad \lambda^{[m]}_1 \geq \cdots \geq \lambda^{[m]}_{D_{m+1}} > 0, \quad \mathrm{tr}\,\Lambda^{[m]} = 1. \tag{4}$$

Notice that

$$\left\langle\psi|\psi\right\rangle = 1. \tag{5}$$

Let $D \in \mathbb{Z}_{>0}$. Introduce the MPS obtained by cutting the bond dimensions after D,

$$\left|\psi_D\right\rangle := \sum_{i_1,\dots,i_N} A^{[1]}_{i_1} P A^{[2]}_{i_2} P \cdots P A^{[N-1]}_{i_{N-1}} P A^{[N]}_{i_N} \left|i_1 \cdots i_N\right\rangle, \quad P = \sum_{k=1}^{D} \left|k\right\rangle \left\langle k\right|. \tag{6}$$

On bonds whose bond dimension is already less than D, P is understood to act trivially. Define the "local error" at the bond between m and $m+1$ by

$$\epsilon_m(D) := \sum_{a>D} \lambda^{[m]}_a. \tag{7}$$

(Lemma 1 in [1])

$$\|\,\left|\psi\right\rangle - \left|\psi_D\right\rangle\|^2 \leq 2 \sum_{m=1}^{N-1} \epsilon_m(D). \tag{8}$$

Proof. It is enough to estimate $|\left\langle\psi_D|\psi\right\rangle|$. We have

$$\left\langle\psi_D|\psi\right\rangle = \mathrm{Tr}\left[S_N\big(PS_{N-1}(PS_{N-2}(\cdots PS_3(PS_2(P\Lambda^{[1]}))\cdots))\big)\right]. \tag{9}$$

Here S_m is the transfer matrix

$$S_m(X) := \sum_i A_i^{[m]\dagger} X A_i^{[m]}, \quad S_{m+1}(\Lambda^{[m]}) = \Lambda^{[m+1]}. \tag{10}$$

It is a positive linear map and is trace preserving:

$$\mathrm{tr}\,[S_m(X)] = \mathrm{tr}\,[X]. \tag{11}$$

Therefore, by contractivity of the trace norm,

$$\|S_m(X)\|_{\mathrm{tr}} \le \|X\|_{\mathrm{tr}}. \tag{12}$$

Introduce the notation

$$Y^{[1]} = \Lambda^{[1]}, \quad Y^{[k+1]} = S_{k+1}(PY^{[k]}) = \sum_i A_i^{[k+1]\dagger} P Y^{[k]} A_i^{[k+1]}. \tag{13}$$

Then

$$\langle \psi_D | \psi \rangle = S_N\big(PS_{N-1}(PS_{N-2}(\cdots PS_3(PS_2(P\Lambda^{[1]}))\cdots))\big) = Y^{[N]}. \tag{14}$$

Also note that

$$S_m(\Lambda^{[m-1]}) = \Lambda^{[m]}. \tag{15}$$

Now

$$\begin{aligned} \| \, |\psi\rangle - |\psi_D\rangle \, \|^2 &= 1 + \langle \psi_D | \psi_D \rangle - 2\Re \langle \psi_D | \psi \rangle \\ &\le 2(1 - \Re Y^{[N]}) \le 2|1 - Y^{[N]}|. \end{aligned} \tag{16}$$

By contractivity of the trace norm,

$$\begin{aligned} |1 - Y^{[N]}| &= |\Lambda^{[N]} - Y^{[N]}| = \|\Lambda^{[N]} - Y^{[N]}\|_{\mathrm{tr}} \\ &= \|S_N(\Lambda^{[N-1]} - PY^{[N-1]})\|_{\mathrm{tr}} \\ &\le \|\Lambda^{[N-1]} - PY^{[N-1]}\|_{\mathrm{tr}} \\ &= \|\Lambda^{[N-1]} - P\Lambda^{[N-1]} + P\Lambda^{[N-1]} - PY^{[N-1]}\|_{\mathrm{tr}} \\ &\le \|\Lambda^{[N-1]} - P\Lambda^{[N-1]}\|_{\mathrm{tr}} + \|P(\Lambda^{[N-1]} - Y^{[N-1]})\|_{\mathrm{tr}}. \end{aligned} \tag{17}$$

Using $\|AB\|_{\mathrm{tr}} \le \|A\|\|B\|_{\mathrm{tr}}$,[1] we obtain

$$|1 - Y^{[N]}| \le \|\Lambda^{[N-1]} - P\Lambda^{[N-1]}\|_{\mathrm{tr}} + \|\Lambda^{[N-1]} - Y^{[N-1]}\|_{\mathrm{tr}}. \tag{18}$$

Repeating the same estimate for $\|\Lambda^{[N-1]} - Y^{[N-1]}\|_{\mathrm{tr}}$, we finally get

$$|1 - Y^{[N]}| \le \sum_{m=1}^{N-1} \|\Lambda^{[m]} - P\Lambda^{[m]}\|_{\mathrm{tr}} = \sum_{m=1}^{N-1} \sum_{a>D} \lambda_a^{[m]}. \tag{19}$$

□

Thus the error from the MPS can be bounded from above by the sum of the errors over all bonds.

Next I prove a general inequality between Renyi entropy and the density matrix. For a density matrix ρ, the Renyi entropy is defined by

$$S^\alpha(\rho) = \frac{1}{1-\alpha} \log \mathrm{tr}\, \rho^\alpha. \tag{20}$$

---

[1] For singular values, one generally has $\sigma_k(AB) \le \sigma_1(A)\sigma_k(B)$. Therefore $\|AB\|_{\mathrm{tr}} = \sum_i \sigma_i(AB) \le \sum_i \sigma_1(A)\sigma_i(B) = \|A\|\|B\|_{\mathrm{tr}}$.

Writing the spectral decomposition as

$$\rho = \sum_a \lambda_a \left|a\right\rangle \left\langle a\right|, \quad \lambda_1 \geq \lambda_2 \geq \cdots, \quad \sum_a \lambda_a = 1, \tag{21}$$

one has

$$S^\alpha(\rho) = \frac{1}{1-\alpha} \log \sum_a \lambda_a^\alpha. \tag{22}$$

Define

$$\epsilon(D) := \sum_{a>D} \lambda_a. \tag{23}$$

(Lemma 2 in [1])[a]

$$\log \frac{\epsilon(D)}{\alpha} \leq \frac{1-\alpha}{\alpha} \left(S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right), \quad 0 < \alpha < 1. \tag{25}$$

[a] In [1] the formula is written as

$$\log \epsilon(D) \leq \frac{1-\alpha}{\alpha} \left(S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right), \quad 0 < \alpha < 1. \tag{24}$$

Is this a typo?

Proof.[2] Let $0 < \alpha < 1$. We estimate $S^\alpha(\rho)$ from below while fixing $\epsilon(D)$. Since log is monotone increasing, it is enough to estimate $\sum_a \lambda_a^\alpha$ from below. The function

$$\left\{ \{p_a\}_{a=1}^N \in [0,1]^{\times N} \,\middle|\, \sum_{a=1}^N p_a = 1 \right\} \to \sum_{a=1}^N p_a^\alpha \tag{26}$$

is symmetric in its variables and convex upward, hence Schur-concave. In other words, it can be bounded from below using a distribution $\{p_a\}_a$ that majorizes $\{\lambda_a\}_a$:[3]

$$\{p_a\}_a \succ \{\lambda_a\}_a \quad \Rightarrow \quad \sum_a p_a^\alpha \leq \sum_a \lambda_a^\alpha. \tag{27}$$

Among all distributions, the distribution $p_1 = 1$, $p_{a>1} = 0$ majorizes any distribution, but it only gives the trivial estimate

$$1 \leq \sum_a \lambda_a^\alpha. \tag{28}$$

Therefore, while fixing the tail

$$\sum_{a>D} p_a = \epsilon(D), \tag{29}$$

we look for the most biased distribution that still majorizes $\{\lambda_a\}_a$.

---

[2] This proof is based on discussions with Akifumi Shimomura about the proof in [1].

[3] Let $p_1^\downarrow \geq p_2^\downarrow \geq \cdots$ be the distribution $\{p_a\}_a$ arranged in descending order. The definition of $\{p_a\}_a \succ \{q_a\}_a$ is that $\sum_{a=1}^k p_a^\downarrow \geq \sum_{a=1}^k q_a^\downarrow$ for every k.

Let $0 < h \leq (1-\epsilon(D))/D$ be a parameter and consider the following distribution $\{q_a\}_a$:

$$
\begin{aligned}
&q_1 = 1-\epsilon(D)-(D-1)h, && (30)\\
&q_2 = q_3 = \cdots = q_D = q_{D+1} = \cdots = q_{D+\lfloor \epsilon(D)/h \rfloor} = h, && (31)\\
&q_{D+\lfloor \epsilon(D)/h \rfloor+1} = \epsilon(D) - \left\lfloor \frac{\epsilon(D)}{h} \right\rfloor h, && (32)\\
&q_{a > D+\lfloor \epsilon(D)/h \rfloor+1} = 0, && (33)\\
&\sum_{a>D} q_a = \left\lfloor \frac{\epsilon(D)}{h} \right\rfloor h + \epsilon(D) - \left\lfloor \frac{\epsilon(D)}{h} \right\rfloor h = \epsilon(D). && (34)
\end{aligned}
$$

Keeping the sums

$$
\sum_{a\leq D} q_a = 1-\epsilon(D), \quad \sum_{a>D} q_a = \epsilon(D) \tag{35}
$$

fixed, we vary the height $q_D = h$ of the middle part.

First estimate the sum $\sum_a q_a^\alpha$ for this distribution:

$$
\begin{aligned}
\sum_a q_a^\alpha &= (1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\lfloor \epsilon(D)/h \rfloor)h^\alpha \\
&\quad + \left(\epsilon(D) - \left\lfloor \frac{\epsilon(D)}{h} \right\rfloor h\right)^\alpha.
\end{aligned} \tag{36}
$$

For simplicity, consider only the case $h = \epsilon(D)/r$, $r \in \mathbb{Z}_{>0}$. Using

$$
\begin{aligned}
&(1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\epsilon(D)/h)h^\alpha - Dh^\alpha - \epsilon(D)h^{\alpha-1} \\
&= (1-\epsilon(D)-(D-1)h)^\alpha - h^\alpha \\
&= (1-\epsilon(D)-Dh)(\cdots) \geq 0,
\end{aligned} \tag{37}
$$

we get

$$
\begin{aligned}
\sum_a q_a^\alpha &= (1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\epsilon(D)/h)h^\alpha \\
&\geq Dh^\alpha + \epsilon(D)h^{\alpha-1}.
\end{aligned} \tag{38}
$$

Since $0 < \alpha < 1$, there is an h-independent lower bound. It is attained at

$$
h = h_* = \frac{(1-\alpha)\epsilon(D)}{\alpha D}. \tag{39}
$$

Thus

$$
\begin{aligned}
Dh^\alpha + \epsilon(D)h^{\alpha-1} &\geq \left(Dh^\alpha + \epsilon(D)h^{\alpha-1}\right)\Big|_{h=\frac{(1-\alpha)\epsilon(D)}{\alpha D}} \\
&= \frac{D\left(\frac{\epsilon(D)(1-\alpha)}{D\alpha}\right)^\alpha}{1-\alpha} = \frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha}.
\end{aligned} \tag{40}
$$

Therefore, if there exists an h satisfying $0 < h \leq (1-\epsilon(D))/D$ and

$$
\{q_a\}_a \succ \{\lambda_a\}_a, \tag{41}
$$

then $\sum_a \lambda_a^\alpha$ is bounded from below by the expression above. The issue is whether such an h can always be found.

For $h = \lambda_D$, one has

$$
\{q_a\}_a \succ \{\lambda_a\}_a. \tag{42}
$$

Indeed, since the area up to $a = 1, \dots, D$ is fixed,

$$\sum_{a=1}^{D} q_a = \sum_{a=1}^{D} \lambda_a = 1 - \epsilon(D), \tag{43}$$

the distribution $\{q_a\}_a$ puts as much weight as possible into q_1, and for $k = 1, \dots, D$,

$$\begin{aligned} \sum_{a=1}^{k} q_a &= 1 - \epsilon(D) - (D-1)h + (k-1)h \\ &= \sum_{a=1}^{D} \lambda_a - (D-k)\lambda_D \\ &= \sum_{a=1}^{k} \lambda_a + \sum_{a=k+1}^{D} (\lambda_a - \lambda_D) \geq \sum_{a=1}^{k} \lambda_a. \end{aligned} \tag{44}$$

For $k > D$, the tail part of λ_a has been made fatter. Hence, for $D < k \leq D + \lfloor \epsilon(D)/h \rfloor$,

$$\begin{aligned} \sum_{a=1}^{k} q_a &= 1 - \epsilon(D) + (k-D)\lambda_D \\ &= \sum_{a=1}^{D} \lambda_a + \sum_{a=D+1}^{k} \lambda_D \geq \sum_{a=1}^{D} \lambda_a + \sum_{a=D+1}^{k} \lambda_a, \end{aligned} \tag{45}$$

and for $k > D + \lfloor \epsilon(D)/h \rfloor$,

$$\sum_{a=1}^{k} q_a = 1 \geq \sum_{a=1}^{k} \lambda_a. \tag{46}$$

Therefore

$$\{q_a\}_a \big|_{h=\lambda_D} \succ \{\lambda_a\}_a. \tag{47}$$

It follows that

$$\frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha} \leq \sum_a \lambda_a^\alpha. \tag{48}$$

Thus we have obtained a lower bound determined only by D and $\epsilon(D)$. Taking the Renyi entropy gives

$$\frac{1}{1-\alpha} \log \frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha} \leq S^\alpha(\rho), \tag{49}$$

that is,

$$\log \frac{D}{1-\alpha} + \frac{\alpha}{1-\alpha} \log \frac{\epsilon(D)}{\alpha} \leq S^\alpha(\rho). \tag{50}$$

Therefore

$$\log \frac{\epsilon(D)}{\alpha} \leq \frac{1-\alpha}{\alpha} \left[S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right], \tag{51}$$

or equivalently

$$\epsilon(D) \leq \alpha \exp\left(\frac{1-\alpha}{\alpha} \left[S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right] \right). \tag{52}$$

□

For a critical system, the Renyi entropy on an interval $[1, L]$ is apparently known to behave as

$$S^\alpha(\rho_L) = \frac{c + \bar{c}}{12} \left(1 + \frac{1}{\alpha} \right) \log L. \tag{53}$$

Therefore, even for a critical system, $\sum_{k=1}^{N-1} \epsilon_k(D)$ is bounded by a power of N.

References

[1] F. Verstraete and J. I. Cirac, *Matrix product states represent ground states faithfully*, arXiv:cond-mat/0505140.