Note: Matrix Product State *

Ken Shiozaki

November 12, 2024

Abstract

The purpose of this note is to provide a proof of the so-called fundamental theorem of matrix product states, which concerns the gauge degrees of freedom of translationally invariant matrix product states. In connection with this, we also follow the proofs of certain facts regarding matrix product states and the accuracy of approximations.

Contents

*This English translation was generated from the original Japanese notes using AI translation by ChatGPT (GPT-5).

Introduction

This note fills in the details between the lines of Refs. [2, 7] concerning the proof of a certain gauge indeterminacy, called the fundamental theorem of matrix product states (MPS) [2][1].

For general aspects of MPS, the review articles [3, 4] are useful references. For matrix analysis, the lecture notes [5] are also helpful. It should be noted that many important results concerning MPS were obtained in Ref. [1]. On the other hand, Ref. [1] is based on a mathematically rigorous framework for infinite systems, and is technically demanding. Ref. [2] extends Ref. [1] to systems with a finite number of sites, and this note follows the proofs related to the fundamental theorem of MPS in Ref. [2].

A matrix product state is a type of tensor network representation of a wavefunction, and in particular in one spatial dimension it is expressed as

$$\psi_{i_1\dots i_N} = A^{[1]}_{i_1} A^{[2]}_{i_2} \cdots A^{[N]}_{i_N}, \tag{0.1}$$

where the wavefunction is written as a product of matrices $A^{[m]}_{i_m}$. For a finite N-site system, there always exists such a set of matrices $\{\{A^{[m]}_{i_m}\}_{i_m}\}_m$ that exactly represents a given wavefunction $\psi_{i_1\dots i_N}$ without any loss of information. The crucial point, however, is that for certain classes of states, for example ground states of Hamiltonians consisting of local terms, depending on the physical quantities of interest, one can efficiently approximate the matrices $A^{[m]}_{i_m}$ by finite-rank approximations, with computational cost scaling only polynomially in the system size N and in the inverse of the error. Consequently, one can investigate the properties of quantum states in one spatial dimension—which are essentially infinite-dimensional systems—using finite-size matrices. As an example, we will see in the final section that, owing to the fundamental theorem of MPS, one can define invariants characterizing symmetry-protected topological (SPT) phases.

1 Approximation of Wavefunctions

1.1 Schmidt Decomposition

We begin with the singular value decomposition.

Theorem 1.1 (Singular Value Decomposition)**.** *For a matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{C})$, the following decomposition holds:*

$$A = U\Sigma V^\dagger, \quad U \in \mathrm{Mat}_{m,r}(\mathbb{C}), \quad V \in \mathrm{Mat}_{n,r}(\mathbb{C}), \quad \Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_r), \tag{1.1}$$

$$U^\dagger U = 1_r, \quad V^\dagger V = 1_r, \quad \sigma_1 \geq \cdots \geq \sigma_r. \tag{1.2}$$

Here $r = \mathrm{rank}(A)$, the set of singular values $\{\sigma_a\}_{a=1}^r$ is uniquely determined, while U and V are unique up to unitary transformations within degenerate singular values.

(Proof) Note that the matrix $A^\dagger A$ is Hermitian and positive semidefinite. Diagonalizing $A^\dagger A$, we obtain

$$A^\dagger A = \sum_{a=1}^r \sigma_a^2 v_a v_a^\dagger, \quad \sigma_1 \geq \cdots \geq \sigma_r > 0, \quad v_a^\dagger v_b = \delta_{ab}. \tag{1.3}$$

[1]The content of this note is based on the intensive lectures "Gauge degrees of freedom of matrix product states and their applications," held at the University of Tokyo on July 29 (Mon.) – July 30 (Tue.), 2024. In preparing this note, I am deeply indebted to Kenji Shimomura, Kan Kitamura, and Akihiro Hokkyo for instructive discussions on numerous questions, as well as for pointing out errors.

Let $V = (v_1, \dots, v_r)$. Defining $u_a = Av_a/\sigma_a,\ a = 1, \dots, r$, we have $u_a^\dagger u_b = v_a^\dagger A^\dagger A v_b/(\sigma_a\sigma_b) = \delta_{ab}$, and

$$\sum_{a=1}^{r} \sigma_a u_a v_a^\dagger = A \sum_{a=1}^{r} v_a v_a^\dagger = A - A(1_n - \sum_{a=1}^{r} v_a v_a^\dagger) = A. \tag{1.4}$$

Here we used $\ker A = \ker A^\dagger A$, which follows from $Av = 0 \Rightarrow A^\dagger Av = 0$ and $A^\dagger Av = 0 \Rightarrow \|Av\| = 0 \Rightarrow Av = 0$. The uniqueness of $\{\sigma_a\}_{a=1}^r$ is clear from the construction. The non-uniqueness of singular vectors v_a arises when singular values are degenerate: for $\sigma_{a_1} = \cdots = \sigma_{a_p}$, the ambiguity is given by a unitary transformation $W \in U(p)$,

$$(v_{a_1}, \dots, v_{a_p}) \mapsto (v_{a_1}, \dots, v_{a_p})W, \tag{1.5}$$

and correspondingly,

$$(u_{a_1}, \dots, u_{a_p}) \mapsto (u_{a_1}, \dots, u_{a_p})W. \tag{1.6}$$

□

We now move to the Schmidt decomposition. Consider a quantum state $|\psi\rangle$ in a general finite-dimensional Hilbert space $\mathcal{H}$,

$$|\psi\rangle = \sum_i \psi_i |i\rangle. \tag{1.7}$$

Here $|i\rangle$ is a basis of the Hilbert space $\mathcal{H}$. Decompose $\mathcal{H}$ into two subsystems L and R in an arbitrary manner:

$$\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R. \tag{1.8}$$

Accordingly, the basis can be written as $|i\rangle = |i_L, i_R\rangle = |i_L\rangle \otimes |i_R\rangle$, and

$$|\psi\rangle = \sum_{i_L, i_R} \psi_{i_L, i_R} |i_L, i_R\rangle. \tag{1.9}$$

Applying the singular value decomposition to the wavefunction ψ_{i_L,i_R}, we obtain the following decomposition.

Corollary 1.2 (Schmidt Decomposition)**.** *For any state $|\psi\rangle \in \mathcal{H}_L \otimes \mathcal{H}_R$, there exists a decomposition*

$$|\psi\rangle = \sum_{a=1}^{r} \lambda_a^{\frac{1}{2}} |L_a\rangle \otimes |R_a\rangle. \tag{1.10}$$

Here $\{|L_a\rangle\}_a$ and $\{|R_a\rangle\}_a$ are orthonormal sets satisfying $\langle L_a|L_b\rangle = \delta_{ab}, \langle R_a|R_b\rangle = \delta_{ab}$, and $\{\lambda_a\}_{a=1}^r$ are positive real numbers with $\lambda_1 \geq \cdots \geq \lambda_r > 0$. The integer $r \in \mathbb{N}$ and the set $\{\lambda_a\}_{a=1}^r$ are uniquely determined. The orthonormal sets $\{L_a\}_{a=1}^r, \{R_a\}_{a=1}^r$ are unique up to unitary transformations within blocks of degenerate Schmidt coefficients λ_a. That is, for a block with $\lambda_{a_1} = \cdots = \lambda_{a_p}$, the ambiguity is given by a unitary $U \in U(p)$:

$$(L_{a_1}, \dots, L_{a_p}) \mapsto (L_{a_1}, \dots, L_{a_p})U, \quad (R_{a_1}, \dots, R_{a_p}) \mapsto (R_{a_1}, \dots, R_{a_p})U^*. \tag{1.11}$$

Moreover, when $|\psi\rangle$ satisfies the normalization condition $\langle\psi|\psi\rangle = 1$, we have $\sum_{a=1}^r \lambda_a = 1$.

- r is called the Schmidt rank.
- The reduced density matrix obtained by tracing out subsystem L is

$$\rho_R = \mathrm{tr}_L |\psi\rangle\langle\psi| = \sum_{a=1}^{r} \lambda_a |R_a\rangle\langle R_a|, \tag{1.12}$$

and similarly,

$$\rho_L = \text{tr}_R \,|\psi\rangle\langle\psi| = \sum_{a=1}^{r} \lambda_a \,|L_a\rangle\langle L_a| \,. \tag{1.13}$$

- In particular, the entanglement entropy is given by

$$S(\rho_R) = -\text{tr}\, \rho_R \log \rho_R = -\sum_{a=1}^{r} \lambda_a \log \lambda_a. \tag{1.14}$$

- Example. $|\uparrow\uparrow\rangle = |\uparrow\rangle \otimes |\uparrow\rangle$ has a Schmidt decomposition with $r = 1, \lambda = 1$.
- Example. The Bell pair

$$\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} |\uparrow\rangle \otimes |\downarrow\rangle + \frac{1}{\sqrt{2}} (-|\downarrow\rangle) \otimes |\uparrow\rangle \tag{1.15}$$

has a Schmidt decomposition with $r = 2, \lambda_1 = \lambda_2 = 1/\sqrt{2}$.

1.2 Low-rank Approximation

The Schmidt decomposition provides the best approximation of a state $|\psi\rangle$ in the following sense.

Given the Schmidt decomposition of the state $|\psi\rangle$,

$$|\psi\rangle = \sum_{a=1}^{r} \lambda_a^{\frac{1}{2}} |L_a\rangle \otimes |R_a\rangle \,, \tag{1.16}$$

we introduce the rank-D approximation

$$|\psi_D\rangle := \sum_{a=1}^{D} \lambda_a^{\frac{1}{2}} |L_a\rangle \otimes |R_a\rangle \,. \tag{1.17}$$

Note that $\langle\psi_D|\psi_D\rangle \leq 1$.

Theorem 1.3. *For any state $|\phi_D\rangle$ of Schmidt rank D, we have*

$$\| \,|\psi\rangle - |\phi_D\rangle\, \| \geq \| \,|\psi\rangle - |\psi_D\rangle\, \| = \sqrt{\sum_{a>D} \lambda_a}. \tag{1.18}$$

Furthermore, when $|\phi_D\rangle$ is normalized, the following holds:

$$|\,\langle\psi|\phi_D\rangle\,| \leq |\,\langle\psi|\psi_D\rangle\,| = \sum_{a=1}^{D} \lambda_a. \tag{1.19}$$

Equation (1.19) follows immediately from (1.18): from $\| \,|\psi\rangle - |\phi_D\rangle\, \|^2 \geq \| \,|\psi\rangle - |\psi_D\rangle\, \|^2$, we obtain

$$2 - 2\Re \langle\psi|\phi_D\rangle \geq 1 + \langle\psi_D|\psi_D\rangle - 2\Re \langle\psi|\psi_D\rangle \geq 2 - 2\Re \langle\psi|\psi_D\rangle \,, \tag{1.20}$$

which implies

$$\Re \langle\psi|\phi_D\rangle \leq \Re \langle\psi|\psi_D\rangle = \sum_{a=1}^{D} \lambda_a. \tag{1.21}$$

The phase of $|\phi_D\rangle$ is arbitrary. In particular, by choosing the phase such that $\langle\psi|\phi_D\rangle \geq 0$, we obtain (1.18).

The corresponding theorem for the singular value decomposition is known as the Eckart–Young–Mirsky theorem. We now proceed to prove this. As preparation, we present the min-max theorem and Weyl's inequality.

Theorem 1.4 (Min-max Theorem)**.** *Let A be an $n \times n$ Hermitian matrix. Denote its eigenvalues by*

$$\lambda_1 \geq \cdots \geq \lambda_n. \tag{1.22}$$

Then,

$$\lambda_k = \max_{\dim V = k} \min_{v \in V, \|v\|=1} \langle v|A|v\rangle , \tag{1.23}$$
$$= \min_{\dim V = n-k+1} \max_{v \in V, \|v\|=1} \langle v|A|v\rangle . \tag{1.24}$$

Clearly,

$$\lambda_1 = \max_{v, \|v\|=1} \langle v|A|v\rangle , \tag{1.25}$$
$$\lambda_n = \min_{v, \|v\|=1} \langle v|A|v\rangle , \tag{1.26}$$

hold. The min-max theorem provides such a variational characterization for the other eigenvalues as well.

(Proof) Let the degeneracy of the eigenvalue λ_a be d_a. List the eigenvalues of A without multiplicities as

$$\lambda_1 > \cdots > \lambda_p, \tag{1.27}$$

and denote the eigenspace corresponding to $\lambda = \lambda_a$ by V_a (note that $\sum_{a=1}^{p} \dim V_a = n$). For a subspace $V \subset \mathbb{C}^n$, the following holds:

$$V \cap (V_a \oplus \cdots \oplus V_p) \neq \{0\} \quad \Rightarrow \quad \min_{v \in V, \|v\|=1} \langle v|A|v\rangle \leq \lambda_a, \tag{1.28}$$
$$V \subset (V_a \oplus \cdots \oplus V_p)^{\perp} \quad \Rightarrow \quad \min_{v \in V, \|v\|=1} \langle v|A|v\rangle > \lambda_a. \tag{1.29}$$

The latter does not hold if $\dim V > \sum_{b=1}^{a-1} \dim V_b$. Moreover, when $\sum_{b=1}^{a-1} \dim V_b < \dim V \leq \sum_{b=1}^{a} \dim V_b$, equality in the latter is attained in the case $V = V_1 \oplus \cdots \oplus V_{a-1} \oplus V'$, where $\{0\} \subsetneq V' \subset V_a$. This yields (1.23). Equation (1.24) is obtained by applying (1.23) to $-A$. □

One can also derive a similar variational formula for singular values. List the singular values of a matrix $A \in \mathrm{Mat}_{n,m}(\mathbb{C})$ in decreasing order as

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)}. \tag{1.30}$$

The squares of the singular values of A are nothing but the eigenvalues of the Hermitian matrix $A^\dagger A$. Noting the trivial inequality

$$\sqrt{\langle v|A^\dagger A|v\rangle} \geq \sqrt{\langle v'|A^\dagger A|v'\rangle} \Leftrightarrow \langle v|A^\dagger A|v\rangle \geq \langle v'|A^\dagger A|v'\rangle , \tag{1.31}$$

we obtain the expression

$$\sigma_k = \max_{\dim V = k} \min_{v \in V, \|v\|=1} \sqrt{\langle v|A^\dagger A|v\rangle} = \max_{\dim V = k} \min_{v \in V, \|v\|=1} \|Av\|. \tag{1.32}$$

Corollary 1.5 (Min-max Theorem for Singular Values)**.** *Let the singular values of an $n \times m$ matrix A be ordered as*

$$\sigma_1 \geq \cdots \geq \sigma_{\min(m,n)}. \tag{1.33}$$

Then,

$$\sigma_k = \max_{\dim V = k} \min_{v \in V, \|v\|=1} \|Av\| \tag{1.34}$$

$$= \min_{\dim V = n-k+1} \max_{v \in V, \|v\|=1} \|Av\|. \tag{1.35}$$

Since we will use it later, we state the following.

Corollary 1.6.

$$\sigma_k(AB) \leq \sigma_1(A)\sigma_k(B). \tag{1.36}$$

(Proof)

$$\sigma_k(AB) = \max_{\dim V = k} \min_{v \in V, \|v\|=1} \|ABv\| \leq \|A\| \max_{\dim V = k} \min_{v \in V, \|v\|=1} \|Bv\| = \sigma_1(A)\sigma_k(B). \quad \square \tag{1.37}$$

Next, we prove Weyl's inequality. Let A be an $n \times n$ Hermitian matrix. List the eigenvalues of A in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n. \tag{1.38}$$

For Hermitian matrices A, B, we want to find inequalities relating the eigenvalue set $\{\lambda_i(A+B)\}_i$ with $\{\lambda_i(A)\}_i$ and $\{\lambda_i(B)\}_i$. For example, what can be said about the maximum eigenvalue? Let v_i^A denote the normalized eigenstate of A corresponding to the eigenvalue $\lambda = \lambda_i(A)$. Then,

$$\lambda_1(A+B) = (v_1^{A+B}, (A+B)v_1^{A+B}) = (v_1^{A+B}, Av_1^{A+B}) + (v_1^{A+B}, Bv_1^{A+B}) \tag{1.39}$$

$$\leq (v_1^A, Av_1^A) + (v_1^B, Bv_1^B) = \lambda_1(A) + \lambda_1(B), \tag{1.40}$$

which shows a valid relation.

More generally, the following holds.

Theorem 1.7 (Weyl's Inequality)**.**

$$\lambda_{i+j-1}(A+B) \leq \lambda_i(A) + \lambda_j(B) \leq \lambda_{i+j-n}(A+B). \tag{1.41}$$

(Proof[2]) By the min-max theorem,

$$\lambda_{i+j-1}(A+B) = \max_{V, \dim V = i+j-1} \min_{x \in V, \|x\|=1} (x, (A+B)x), \tag{1.42}$$

$$\lambda_i(A) = \min_{V, \dim V = n-i+1} \max_{x \in V, \|x\|=1} (x, Ax), \tag{1.43}$$

$$\lambda_j(B) = \min_{V, \dim V = n-j+1} \max_{x \in V, \|x\|=1} (x, Bx). \tag{1.44}$$

[2] Wikipedia, *Weyl's inequality*, url.

Thus there exist subspaces V_{A+B}, V_A, V_B such that

$$\lambda_{i+j-1}(A+B) = \min_{x\in V_{A+B},\|x\|=1}(x,(A+B)x), \quad \dim V_{A+B} = i+j-1, \tag{1.45}$$

$$\lambda_i(A) = \max_{x\in V_A,\|x\|=1}(x,Ax), \quad \dim V_A = n-i+1, \tag{1.46}$$

$$\lambda_j(B) = \max_{x\in V_B,\|x\|=1}(x,Bx), \quad \dim V_B = n-j+1. \tag{1.47}$$

In this case,

$$\dim(V_A\cap V_B) = \dim V_A + \dim V_B - \dim(V_A\cup V_B) \geq \dim V_A + \dim V_B - n = n-i-j+2, \tag{1.48}$$

and

$$\dim(V_{A+B}\cap(V_A\cap V_B)) = \dim V_{A+B} + \dim(V_A\cap V_B) - \dim(V_{A+B}\cup(V_A\cap V_B)) \tag{1.49}$$

$$\geq (i+j-1)+(n-i-j+2)-n = 1, \tag{1.50}$$

so one can take a nonzero vector

$$x_0 \in V_{A+B}\cap V_A\cap V_B. \tag{1.51}$$

Then,

$$\lambda_{i+j-1}(A+B) \leq (x_0,(A+B)x_0) = (x_0,Ax_0)+(x_0,Bx_0) \leq \lambda_i(A)+\lambda_j(B), \tag{1.52}$$

which gives the left-hand side of the inequality. The right-hand side of the claim is obtained by applying the left-hand inequality to (-1) times the matrices. □

As a corollary, we obtain the following.

Corollary 1.8 (Weyl's Inequality for Singular Values)**.**

$$\sigma_{i+j-1}(A+B) \leq \sigma_i(A)+\sigma_j(B). \tag{1.53}$$

(Proof) The singular values of A are the nonnegative eigenvalues of the Hermitian matrix

$$\tilde{A} = \begin{pmatrix} & A \\ A^\dagger & \end{pmatrix}. \tag{1.54}$$

Defining

$$\lambda_i = \sigma_i \quad i=1,\dots,n, \quad \lambda_{n+1} = -\sigma_{n-i+1} \quad i=1,\dots,n, \tag{1.55}$$

Weyl's inequality

$$\lambda_{i+j-1}(\tilde{A}+\tilde{B}) \leq \lambda_i(\tilde{A})+\lambda_j(\tilde{B}) \leq \lambda_{i+j-2n}(\tilde{A}+\tilde{B}) \tag{1.56}$$

holds. Restricting to $i+j-1\leq n$ gives the desired result. □

Next, we prove the Eckart–Young–Mirsky theorem. For an $m\times n$ matrix $A\in \mathrm{Mat}_{m,n}(\mathbb{C})$, let its singular value decomposition be

$$A = U\Sigma V^\dagger = \sum_{i=1}^{\mathrm{rank}(A)} \sigma_i u_i v_i^\dagger, \quad \sigma_1\geq\sigma_2\geq\dots \tag{1.57}$$

The Frobenius norm (defined via the Hilbert–Schmidt inner product) is given by

$$\|A\|_F := \sqrt{(A,A)_{\mathrm{HS}}} = \sqrt{\mathrm{tr}\,[A^\dagger A]} = \sqrt{\sum_{i=1}^{\mathrm{rank}(A)}\sigma_i^2}. \tag{1.58}$$

The 2-norm (operator norm) is defined as

$$\|A\| := \sup_{x\in\mathbb{C}^n, \|x\|=1} \|Ax\| = \sigma_1(A). \tag{1.59}$$

Now, define the rank-k approximation of A by

$$A_k := \sum_{i=1}^{k} \sigma_i u_i v_i^\dagger. \tag{1.60}$$

The claim is the following.

Theorem 1.9 (Eckart–Young–Mirsky)**.** *In both the 2-norm and the Frobenius norm, the best approximation of A by a rank-k matrix is given by A_k. That is, for any rank-k matrix B_k, we have*

$$\|A - B_k\|_F \geq \|A - A_k\|_F, \tag{1.61}$$
$$\|A - B_k\|_2 \geq \|A - A_k\|_2. \tag{1.62}$$

(Proof[3]) By Weyl's inequality, for any i, j,

$$\sigma_{i+j-1}(X+Y) \leq \sigma_i(X) + \sigma_j(Y). \tag{1.63}$$

Let the rank of B be k. Then note that $\sigma_{i>k}(B) = 0$. Taking $j = k+1, X = A - B, Y = B$, we obtain

$$\sigma_{i+k}(A) \leq \sigma_i(A-B) + \sigma_{k+1}(B) = \sigma_i(A-B). \tag{1.64}$$

Therefore,

$$\|A - B\|_2 = \sigma_1(A-B) \geq \sigma_{k+1}(A) = \sigma_1(A - A_k). \tag{1.65}$$

$$\|A - B\|_F^2 = \sum_i \sigma_i^2(A-B) \geq \sum_i \sigma_{i+k}^2(A) = \|A - A_k\|_F^2. \quad \square \tag{1.66}$$

Note that when the wavefunction ψ_{i_L,i_R} is regarded as a matrix, its Frobenius norm coincides with the norm of the state $|\psi\rangle$:

$$\|\psi\|_F^2 = \mathrm{tr}\, \psi^\dagger \psi = \sum_{i_L,i_R} \psi^*_{i_L,i_R} \psi_{i_L,i_R} = \sum_{i_L,i_R} |\psi_{i_L,i_R}|^2 = \| \, |\psi\rangle \, \|^2. \tag{1.67}$$

Therefore, the Eckart–Young–Mirsky theorem with respect to the Frobenius norm is nothing but (1.18).

1.3 Error Estimate

How good is the rank-D approximation $|\psi_D\rangle$ obtained from the Schmidt decomposition? We show that the error can be upper-bounded in terms of the Rényi entropy [6]. In the following, we assume that $\mathrm{rank}(\rho)$ is sufficiently large.

For a general density matrix ρ, the Rényi entropy is defined by

$$S^\alpha(\rho) = \frac{1}{1-\alpha} \log \mathrm{tr}\, \rho^\alpha, \quad \alpha > 0. \tag{1.68}$$

[3]Physics Stack Exchange, *Proof of Eckart-Young-Mirsky theorem*, url.

The limit $\alpha \to 1$ gives the von Neumann entropy:

$$S^{\alpha\to 1}(\rho) = S_{\mathrm{vN}}(\rho) = -\mathrm{tr}\,\rho\log\rho. \tag{1.69}$$

With the spectral decomposition

$$\rho = \sum_a \lambda_a |a\rangle\langle a|, \quad \lambda_1 \geq \lambda_2 \geq \cdots, \quad \sum_a \lambda_a = 1, \tag{1.70}$$

we have

$$S^\alpha(\rho) = \frac{1}{1-\alpha}\log\sum_a \lambda_a^\alpha. \tag{1.71}$$

Note that the function $f(\{\lambda_a\}_a) = \sum_a \lambda_a^\alpha$ is concave for $0 < \alpha < 1$ and convex for $\alpha > 1$. Define the error

$$\epsilon(D) := \sum_{a>D} \lambda_a, \tag{1.72}$$

and estimate $\epsilon(D)$. By (1.18), $\epsilon(D)$ is nothing but the squared error $\| |\psi\rangle - |\psi_D\rangle \|^2$ of the rank-D approximation obtained from the Schmidt decomposition of the state $|\psi\rangle$.

Theorem 1.10 (Lemma 2 in [6]).

$$\log\frac{\epsilon(D)}{\alpha} \leq \frac{1-\alpha}{\alpha}\left(S^\alpha(\rho) - \log\frac{D}{1-\alpha}\right), \quad 0 < \alpha < 1. \tag{1.73}$$

Equivalently,

$$\epsilon(D) \leq \alpha \times \left(\frac{1-\alpha}{D}\right)^{\frac{1-\alpha}{\alpha}} \times \exp\left[\frac{1-\alpha}{\alpha}S^\alpha(\rho)\right], \quad 0 < \alpha < 1, \tag{1.74}$$

holds[4]. Note that the Rényi entropy depends only on the density matrix ρ. The error $\epsilon(D)$ is thus bounded from above by a power law in D.

(Proof[5]) Let $0 < \alpha < 1$. We fix $\epsilon(D)$ and derive a lower bound on $S^\alpha(\rho)$. Since log is monotone increasing, it suffices to lower-bound $\sum_a \lambda_a^\alpha$. The function

$$\left\{ \{p_a\}_a^N \in [0,1]^{\times N} \mid \sum_{a=1}^N p_a = 1 \right\} \to \sum_{a=1}^N p_a^\alpha \tag{1.76}$$

is symmetric and concave in its variables for $0 < \alpha < 1$, hence Schur-concave. That is, it is bounded from below by distributions $\{p_a\}_a$ that majorize $\{\lambda_a\}_a$[6]:

$$\{p_a\}_a \succ \{\lambda_a\}_a \Rightarrow \sum_a p_a^\alpha \leq \sum_a \lambda_a^\alpha. \tag{1.77}$$

If we consider all possible distributions, then the extremal distribution $p_1 = 1, p_{a>1} = 0$ majorizes any distribution, but yields only the trivial bound $1 \leq \sum_a \lambda_a^\alpha$. We therefore fix the tail

$$\sum_{a>D} p_a = \epsilon(D) \tag{1.78}$$

[4] In [6] it is written as

$$\log\epsilon(D) \leq \frac{1-\alpha}{\alpha}\left(S^\alpha(\rho) - \log\frac{D}{1-\alpha}\right), \quad 0 < \alpha < 1, \tag{1.75}$$

which appears to be a typo.

[5] Based on discussions with Kenji Shimomura regarding the proof in [6].

[6] Let $p_1^\downarrow \geq p_2^\downarrow \geq \cdots$ denote the decreasing rearrangement of $\{p_a\}_a$. We say $\{p_a\}_a \succ \{q_a\}_a$ if for all k one has $\sum_{a=1}^k p_a^\downarrow \geq \sum_{a=1}^k q_a^\downarrow$.

and search for a distribution that is as skewed as possible while still majorizing $\{\lambda_a\}_a$.

For a parameter $0 < h \le (1-\epsilon(D))/D$, consider the distribution $\{q_a\}_a$ defined by

$$
\begin{aligned}
&q_1 = 1 - \epsilon(D) - (D-1)h, && (1.79)\\
&q_2 = q_3 = \cdots = q_D = q_{D+1} = \cdots = q_{D+\lfloor \epsilon(D)/h \rfloor} = h, && (1.80)\\
&q_{D+\lfloor \epsilon(D)/h \rfloor + 1} = \epsilon(D) - \lfloor \epsilon(D)/h \rfloor h, && (1.81)\\
&q_{a > D+\lfloor \epsilon(D)/h \rfloor + 1} = 0, && (1.82)\\
&\sum_{a>D} q_a = \lfloor \epsilon(D)/h \rfloor h + \epsilon(D) - \lfloor \epsilon(D)/h \rfloor h = \epsilon(D). && (1.83)
\end{aligned}
$$

Thus the partial sums

$$\sum_{a\le D} q_a = 1-\epsilon(D), \quad \sum_{a>D} q_a = \epsilon(D) \tag{1.84}$$

are kept fixed, while the intermediate "height" $q_D = h$ is varied.

First, evaluate $\sum_a q_a^\alpha$:

$$\sum_a q_a^\alpha = (1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\lfloor \epsilon(D)/h \rfloor)h^\alpha + \big(\epsilon(D) - \lfloor \epsilon(D)/h \rfloor h\big)^\alpha. \tag{1.85}$$

For simplicity, restrict to $h = \epsilon(D)/r$ with $r \in \mathbb{Z}_{>0}$. Observing that

$$
\begin{aligned}
&(1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\epsilon(D)/h)h^\alpha - Dh^\alpha - \epsilon(D)h^{\alpha-1} && (1.86)\\
&= (1-\epsilon(D)-(D-1)h)^\alpha - h^\alpha && (1.87)\\
&= (1-\epsilon(D)-Dh)\,(\cdots) \ge 0, && (1.88)
\end{aligned}
$$

we obtain

$$\sum_a q_a^\alpha = (1-\epsilon(D)-(D-1)h)^\alpha + (D-1+\epsilon(D)/h)h^\alpha \ge Dh^\alpha + \epsilon(D)h^{\alpha-1}. \tag{1.89}$$

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

$$h = h_* = \frac{(1-\alpha)\epsilon(D)}{\alpha D}. \tag{1.90}$$

Hence,

$$Dh^\alpha + \epsilon(D)h^{\alpha-1} \ge \Big(Dh^\alpha + \epsilon(D)h^{\alpha-1}\Big)\Big|_{h=\frac{(1-\alpha)\epsilon(D)}{\alpha D}} = \frac{D\big(\frac{\epsilon(D)(1-\alpha)}{D\alpha}\big)^\alpha}{1-\alpha} = \frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha}. \tag{1.91}$$

Therefore, if there exists h with $0 < h \le (1-\epsilon(D))/D$ such that $\{q_a\}_a \succ \{\lambda_a\}_a$, then $\sum_a \lambda_a^\alpha$ is lower-bounded by the right-hand side above. We now address the existence of such an h.

For $h = \lambda_D$, one has $\{q_a\}_a \succ \{\lambda_a\}_a$. Indeed, since the area up to $a = D$ is fixed,

$$\sum_{a=1}^{D} q_a = \sum_{a=1}^{D} \lambda_a = 1-\epsilon(D), \tag{1.92}$$

the distribution $\{q_a\}_a$ that pushes all weight to q_1 satisfies

$$
\begin{aligned}
\sum_{a=1}^{k} q_a &= 1-\epsilon(D)-(D-1)h+(k-1)h = \sum_{a=1}^{D}\lambda_a - (D-k)\lambda_D && (1.93)\\
&= \sum_{a=1}^{k}\lambda_a + \sum_{a=k+1}^{D}(\lambda_a - \lambda_D) \ge \sum_{a=1}^{k}\lambda_a, \quad k = 1,\ldots,D, && (1.94)
\end{aligned}
$$

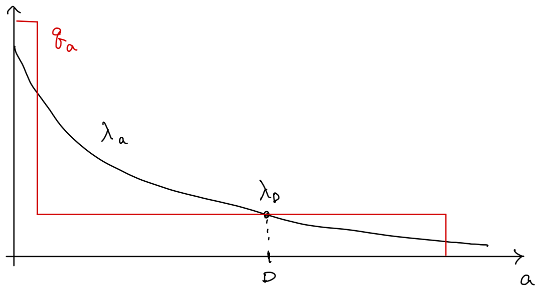

Figure 1

as illustrated in Fig. 1. Furthermore, for $k > D$, because we have fattened the tail of λ_a, for $D < k \leq D + \lfloor \epsilon(D)/h \rfloor$,

$$\sum_{a=1}^{k} \lambda_a = 1 - \epsilon(D) + (k - D)\lambda_D = \sum_{a=1}^{D} \lambda_a + \sum_{a=D+1}^{k} \lambda_D \geq \sum_{a=1}^{D} \lambda_a + \sum_{a=D+1}^{k} \lambda_a, \tag{1.95}$$

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

$$\sum_{a=1}^{k} \lambda_a = 1 \geq \sum_{a=1}^{k} \lambda_a. \tag{1.96}$$

Hence,

$$\{q_a\}_a \big|_{h=\lambda_D} \succ \{\lambda_a\}_a. \tag{1.97}$$

Therefore,

$$\frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha} \leq \sum_a \lambda_a^\alpha. \tag{1.98}$$

Thus we have obtained a lower bound depending only on D and $\epsilon(D)$. Taking the Rényi entropy,

$$\frac{1}{1-\alpha} \log \frac{D^{1-\alpha}\epsilon(D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha} \leq S^\alpha(\rho), \tag{1.99}$$

that is,

$$\log \frac{D}{1-\alpha} + \frac{\alpha}{1-\alpha} \log \frac{\epsilon(D)}{\alpha} \leq S^\alpha(\rho), \tag{1.100}$$

and hence,

$$\log \frac{\epsilon(D)}{\alpha} \leq \frac{1-\alpha}{\alpha} \left[S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right], \tag{1.101}$$

or equivalently,

$$\epsilon(D) \leq \alpha \exp\left(\frac{1-\alpha}{\alpha} \left[S^\alpha(\rho) - \log \frac{D}{1-\alpha} \right] \right). \tag{1.102}$$

□

In critical systems, the Rényi entropy on an interval $[1, L]$ is known to behave as

$$S^\alpha(\rho_L) = \frac{c + \bar{c}}{12} \left(1 + \frac{1}{\alpha} \right) \log L. \tag{1.103}$$

Therefore, even in critical systems, note that $\sum_{k=1}^{N-1} \epsilon_k(D)$ can be bounded by a power of N. We will examine this point in detail in the next section.

2 OBC-MPS

In this section, we introduce matrix product states (MPS) with open boundary conditions (OBC). An important remark is that OBC-MPS can always be defined, regardless of whether the state itself satisfies periodic boundary conditions.

Although using diagrammatic notation improves readability for MPS, we will not employ it in this note.

2.1 Derivation of MPS

Consider a one-dimensional spin system with N sites. Let d be the dimension of the local Hilbert space; the total Hilbert space is the tensor product of local Hilbert spaces:

$$\mathcal{H} = \bigotimes_{x=1}^{N} \mathcal{H}_x, \quad \mathcal{H}_x \cong \mathbb{C}^d. \tag{2.1}$$

Write a basis of $\mathcal{H}_x$ as $\{|i_x\rangle\}_{i_x=1}^{d}$, and abbreviate

$$|i_1 \dots i_N\rangle = |i_1\rangle \otimes \cdots \otimes |i_N\rangle . \tag{2.2}$$

A state in $\mathcal{H}$ can be written as

$$|\psi\rangle = \sum_{i_1,\dots,i_N} \psi_{i_1 \dots i_N} |i_1 \dots i_N\rangle , \tag{2.3}$$

and we assume the normalization

$$\langle\psi|\psi\rangle = 1. \tag{2.4}$$

Definition 2.1. We call the following representation of the wavefunction $\psi_{i_1 \dots i_N}$ an OBC-MPS:

$$\psi_{i_1 \dots i_N} = A^{[1]}_{i_1} A^{[2]}_{i_2} \cdots A^{[N-1]}_{i_{N-1}} A^{[N]}_{i_N}. \tag{2.5}$$

Here each $A^{[x]}_{i_x}$ is a $D_{x-1} \times D_x$ matrix with $D_0 = D_{N+1} = 1$.

An OBC-MPS always exists. Including existence, we establish the existence of the following standard OBC-MPS.

Theorem 2.2 ([2]. Standard form of OBC-MPS)**.** *There exists an OBC-MPS satisfying*

$$\sum_i A_i^{[m]} A_i^{[m]\dagger} = 1_{r_{m-1}}, \quad 1 \le m \le N, \tag{2.6}$$

$$\sum_i A_i^{[m]\dagger} \Lambda^{[m-1]} A_i^{[m]} = \Lambda^{[m]}, \quad 1 \le m \le N. \tag{2.7}$$

Here $\Lambda^{[m]}$ is a diagonal matrix with positive real entries,

$$\Lambda^{[m]} = \mathrm{diag}(\lambda_1^{[m]}, \dots, \lambda_{r_m}^{[m]}), \tag{2.8}$$

and satisfies

$$\Lambda^{[0]} = \Lambda^{[N]} = 1, \tag{2.9}$$

$$\lambda_1^{[m]} \ge \cdots \ge \lambda_{r_m}^{[m]} > 0, \tag{2.10}$$

$$\mathrm{tr}\, \Lambda^{[m]} = \sum_{a=1}^{r_m} \lambda_a^{[m]} = 1. \tag{2.11}$$

Moreover, r_m is the Schmidt rank associated with the bipartition $[1 \cdots m]$ and $[m+1 \cdots N]$.

The construction is obtained by repeatedly performing singular value decompositions from the right; we only sketch the proof.

(Sketch of proof) First apply SVD to $\psi_{i_1 \dots i_{N-1}, i_N}$ to obtain

$$\psi_{i_1 \dots i_{N-1}, i_N} = \sum_{a_{N-1}} \left(\lambda_{a_{N-1}}^{[N-1]}\right)^{\frac{1}{2}} \psi'_{i_1 \cdots i_{N-1} a_{N-1}} \, [A_{i_N}^{[N]}]_{a_{N-1}}. \tag{2.12}$$

Next, apply SVD to

$$\tilde{\psi}'_{i_1 \cdots i_{N-2},\, i_{N-1} a_{N-1}} = \left(\lambda_{a_{N-1}}^{[N-1]}\right)^{\frac{1}{2}} \psi'_{i_1 \cdots i_{N-1} a_{N-1}}, \tag{2.13}$$

to get

$$\tilde{\psi}'_{i_1 \cdots i_{N-2},\, i_{N-1} a_{N-1}} = \sum_{a_{N-2}} \left(\lambda_{a_{N-2}}^{[N-2]}\right)^{\frac{1}{2}} \psi''_{i_1 \cdots i_{N-2} a_{N-2}} \, A_{a_{N-2} a_{N-1}}^{[N-2]}. \tag{2.14}$$

Proceed similarly for the remaining steps. Using $U^\dagger U = 1$ and $V^\dagger V = 1$ from the SVD $A = U \Sigma V^\dagger$, one derives the conditions (2.6) and (2.7). □

With the standard OBC-MPS in hand, the reduced density matrix is obtained immediately:

$$\rho_{[m+1,N]} = \mathrm{tr}_{\,[1,m]} \left|\psi\right\rangle \left\langle\psi\right| = \sum_{a_m=1}^{r_m} \Lambda_{a_m}^{[m]} \left|R_{a_m}\right\rangle \left\langle R_{a_m}\right|. \tag{2.15}$$

Here,

$$\left|R_{a_m}\right\rangle = \sum_{i_{m+1} \cdots i_N} \sum_{a_{m+1} \cdots a_{N-1}} [A_{i_{m+1}}^{[m+1]}]_{a_m a_{m+1}} \cdots [A_{i_N}^{[N]}]_{a_{N-1}} \left|i_{m+1} \cdots i_N\right\rangle. \tag{2.16}$$

Proposition 2.3. *The standard OBC-MPS is unique up to the indeterminacy of singular vectors.*

This is clear from the construction.

The uniqueness of the standard OBC-MPS plays an important role in the proof of gauge indeterminacy of PBC-MPS.

2.2 Transformation to the Standard OBC-MPS

For a quantum state $|\psi\rangle$, consider an OBC-MPS not necessarily in the standard form,

$$\psi_{i_1\cdots i_N} = B_{i_1}^{[1]}\cdots B_{i_N}^{[N]}. \tag{2.17}$$

We describe a procedure to transform it into a standard OBC-MPS.

Theorem 2.4 ([2]). *For the OBC-MPS (2.17), there exist matrices Y_m, Z_m such that*

$$\begin{aligned}
&A_i^{[m]} \text{ is a standard OBC-MPS,} && (2.18)\\
&Y_m Z_m = 1, \quad 1 < m < N, && (2.19)\\
&A_i^{[1]} = B_i^{[1]} Z_1, \quad A_i^{[N]} = Y_{N-1} B_i^{[N]}, && (2.20)\\
&A_i^{[m]} = Y_{m-1} B_i^{[m]} Z_m, \quad 1 < m < N. && (2.21)
\end{aligned}$$

The derivation is obtained simply by repeatedly applying singular value decompositions from the right; we only sketch the proof.

(Sketch of proof) First, apply the singular value decomposition to $[B_{i_N}^{[N]}]_{b_{N-1}}$:

$$[B_{i_N}^{[N]}]_{b_{N-1}} = \sum_{a_{N-1}} U^{[N-1]}_{b_{N-1}a_{N-1}} \Delta^{[N-1]}_{a_{N-1}} A^{[N]}_{i_N}. \tag{2.22}$$

Define

$$Z_{N-1} = U^{[N-1]}\Delta^{[N-1]}, \quad Y_{N-1} = (\Delta^{[N-1]})^{-1} U^{[N-1]\dagger}, \tag{2.23}$$

so that $Y_{N-1}B_i^{[N]} = A_i^{[N]}$ and $Y_{N-1}Z_{N-1} = 1$.

Next, apply the singular value decomposition to the matrix

$$\sum_{b_{N-1}} [B^{[N-2]}_{i_{N-1}}]_{b_{N-2}b_{N-1}} [Z_{N-1}]_{b_{N-1}a_{N-1}}, \tag{2.24}$$

considered as a bipartition between the legs $b_{N-2}i_{N-1}$ and a_{N-1}:

$$\sum_{b_{N-1}} [B^{[N-2]}_{i_{N-1}}]_{b_{N-2}b_{N-1}} [Z_{N-1}]_{b_{N-1}a_{N-1}} = \sum_{a_{N-2}} U^{[N-2]}_{b_{N-2}a_{N-2}} \Delta^{[N-2]}_{a_{N-2}} [A^{[N-1]}_{i_{N-1}}]_{a_{N-2}a_{N-1}}. \tag{2.25}$$

Defining

$$Z_{N-2} = U^{[N-2]}\Delta^{[N-2]}, \quad Y_{N-2} = (\Delta^{[N-2]})^{-1} U^{[N-2]\dagger}, \tag{2.26}$$

we obtain $Y_{N-2}B_i^{[N-1]}Z_{N-1} = A_i^{[N-1]}$ and $Y_{N-2}Z_{N-2} = 1$.

Repeating this process down to $m = 1$, we arrive at

$$B_{i_1}^{[1]}\cdots B_{i_N}^{[N]} = A_{i_1}^{[1]}\cdots A_{i_N}^{[N]}, \tag{2.27}$$

where the $A_i^{[m]}$ satisfy all conditions except the left-canonical one.

To impose the left-canonical condition, start from $m = 1$ and diagonalize

$$\sum_i A_i^{[1]\dagger} A_i^{[1]} = V_1 \Lambda^{[1]} V_1^\dagger. \tag{2.28}$$

Then $A_i'^{[1]} = A_i^{[1]} V_1$ satisfies the left-canonical condition $\sum_i A_i'^{[1]\dagger} A_i'^{[1]} = \Lambda^{[1]}$, noting that $\Lambda^{[0]} = 1$. Redefine $A_i'^{[2]} = V_1^\dagger A_i^{[2]}$, and diagonalize

$$\sum_i A_i'^{[2]\dagger} \Lambda^{[1]} A_i'^{[2]} = V_2 \Lambda^{[2]} V_2^\dagger. \tag{2.29}$$

Then

$$\sum_i (A_i'^{[2]} V_2)^\dagger \Lambda^{[1]} A_i'^{[2]} V_2 = \Lambda^{[2]}, \tag{2.30}$$

so $A_i''^{[2]} = A_i'^{[2]} V_2$ satisfies the left-canonical condition.

Collecting these transformations,

$$A_i'^{[1]} = A^{[1]} V_1 = B^{[1]} Z_1 V_1, \tag{2.31}$$

$$A_i''^{[m]} = V_{m-1}^\dagger A_i^{[m]} V_m = V_{m-1}^\dagger Y_{m-1} B_i^{[m]} Z_m V_m, \quad m > 1, \tag{2.32}$$

which corresponds to the redefinitions

$$Z_m \mapsto Z_m V_m, \quad Y_m \mapsto V_m^\dagger Y_m, \tag{2.33}$$

while preserving $Y_m Z_m = 1$. □

2.3 Error Estimate

We now estimate how well the OBC-MPS obtained by uniformly truncating the bond dimension of a standard OBC-MPS to D approximates the original state. Let

$$\psi_{i_1 \cdots i_N} = A_{i_1}^{[1]} \cdots A_{i_N}^{[N]} \tag{2.34}$$

be a standard OBC-MPS. Choose $D \in \mathbb{N}$ and define the truncated OBC-MPS by discarding indices $a > D$:

$$\psi_{i_1 \cdots i_N}^D := \sum_{a_1, \ldots, a_{N-1}=1}^{D} [A_{i_1}^{[1]}]_{a_1} [A_{i_2}^{[2]}]_{a_1 a_2} \cdots [A_{i_{N-1}}^{[N-1]}]_{a_{N-2} a_{N-1}} [A_{i_N}^{[N]}]_{a_{N-1}}. \tag{2.35}$$

Equivalently, this can be written using the projection matrix onto the degrees of freedom $a_m \leq D$:

$$P = \begin{pmatrix} 1_D & \\ & O \end{pmatrix}, \tag{2.36}$$

as

$$\psi_{i_1 \cdots i_N}^D = A_{i_1}^{[1]} P A_{i_2}^{[2]} P \cdots P A_{i_{N-1}}^{[N-1]} P A_{i_N}^{[N]}. \tag{2.37}$$

Introduce the "error" at the i_m–i_{m+1} bond by

$$\epsilon_m(D) = \sum_{a_m=1}^{r_m} \lambda_a^{[m]}. \tag{2.38}$$

Proposition 2.5 (Lemma 1 in [6]).

$$\| \, |\psi\rangle - |\psi^D\rangle \, \|^2 \leq 2 \sum_{m=1}^{N-1} \epsilon_m(D). \tag{2.39}$$

The proof requires the contractivity of the trace norm under positive maps (Theorem 3.23), whose proof will be given in the next section.

(Proof) It suffices to evaluate $|\langle\psi^D|\psi\rangle|$. We have

$$\langle\psi^D|\psi\rangle = \mathrm{Tr}\left[S_N(PS_{N-1}(PS_{N-2}(\cdots PS_3(PS_2(P\Lambda^{[1]}))\cdots)))\right] \tag{2.40}$$

where S_m is the transfer matrix defined by

$$S_m(X) := \sum_i A_i^{[m]\dagger} X A_i^{[m]}, \quad S_{m+1}(\Lambda^{[m]}) = \Lambda^{[m+1]}, \tag{2.41}$$

which is a positive linear map and trace-preserving: $\mathrm{tr}\,[S_m(X)] = \mathrm{tr}\,[X]$. Thus, by the contractivity of the trace norm (Theorem 3.23),

$$\|S_m(X)\|_{\mathrm{tr}} \le \|X\|_{\mathrm{tr}}. \tag{2.42}$$

Introduce the notation

$$Y^{[1]} = \Lambda^{[1]}, \quad Y^{[k+1]} = S_{k+1}(PY^{[k]}) = \sum_i A_i^{[m]\dagger} P Y^{[k]} A_i^{[m]}, \tag{2.43}$$

so that

$$\langle\psi^D|\psi\rangle = S_N(PS_{N-1}(PS_{N-2}(\cdots PS_3(PS_2(P\Lambda^{[1]}))\cdots))) = Y^{[N]}. \tag{2.44}$$

Also note that

$$S_m(\Lambda^{[m-1]}) = \Lambda^{[m]}. \tag{2.45}$$

Now,

$$\|\,|\psi\rangle - |\psi^D\rangle\,\|^2 = 1 + \langle\psi^D|\psi^D\rangle - 2\Re\,\langle\psi^D|\psi\rangle \le 2(1-\Re Y^{[N]}) \le 2|1-Y^{[N]}|. \tag{2.46}$$

By the contractivity of the trace norm,

$$\begin{aligned}
|1-Y^{[N]}| = |\Lambda^{[N]} - Y^{[N]}| = \|\Lambda^{[N]} - Y^{[N]}\|_{\mathrm{tr}} &= \|S_N(\Lambda^{[N-1]} - PY^{[N-1]})\|_{\mathrm{tr}} &(2.47)\\
&\le \|\Lambda^{[N-1]} - PY^{[N-1]}\|_{\mathrm{tr}} = \|\Lambda^{[N-1]} - P\Lambda^{[N-1]} + P\Lambda^{[N-1]} - PY^{[N-1]}\|_{\mathrm{tr}} &(2.48)\\
&\le \|\Lambda^{[N-1]} - P\Lambda^{[N-1]}\|_{\mathrm{tr}} + \|P(\Lambda^{[N-1]} - Y^{[N-1]})\|_{\mathrm{tr}}. &(2.49)
\end{aligned}$$

Using $\|AB\|_{\mathrm{tr}} \le \|A\|\|B\|_{\mathrm{tr}}$[7], we obtain

$$|1-Y^{[N]}| \le \|\Lambda^{[N-1]} - P\Lambda^{[N-1]}\|_{\mathrm{tr}} + \|\Lambda^{[N-1]} - Y^{[N-1]}\|_{\mathrm{tr}}. \tag{2.50}$$

Proceeding similarly to evaluate $\|\Lambda^{[N-1]} - Y^{[N-1]}\|_{\mathrm{tr}}$, we finally obtain

$$|1-Y^{[N]}| \le \sum_{m=1}^{N-1} \|\Lambda^{[m]} - P\Lambda^{[m]}\|_{\mathrm{tr}} = \sum_{m=1}^{N-1}\sum_{a>D} \lambda_a^{[m]}. \tag{2.51}$$

□

Therefore, the error between the true state $|\psi\rangle$ and the bond-dimension-D approximation $|\psi^D\rangle$ constructed from the standard OBC-MPS can be bounded from above by the sum of the errors at all bonds. Combining with (1.73), we obtain

$$\|\,|\psi\rangle - |\psi^D\rangle\,\|^2 \le 2\sum_{m=1}^{N-1} \epsilon_m(D) \le 2\sum_{m=1}^{N-1} \alpha \times \left(\frac{1-\alpha}{D}\right)^{\frac{1-\alpha}{\alpha}} \times \exp\left[\frac{1-\alpha}{\alpha} S^\alpha(\rho_{1,m})\right]. \tag{2.52}$$

[7] By Lemma 1.6, $\|AB\|_{\mathrm{tr}} = \sum_i \sigma_i(AB) \le \sum_i \sigma_1(A)\sigma_i(B) = \|A\|\|B\|_{\mathrm{tr}}$.

Now, when $|\psi\rangle$ is the ground state of a CFT, for $1 \ll \ell \ll N$ we have

$$S^{\alpha}(\rho_{1,\ell}) \sim \frac{c}{6}\Big(1+\frac{1}{\alpha}\Big)\log \ell. \tag{2.53}$$

Hence,

$$\| \, |\psi\rangle - |\psi^D\rangle \, \|^2 \le (\text{const.}) \times D^{-\frac{1-\alpha}{\alpha}} N^{\frac{c}{6}(1+\frac{1}{\alpha})+1}. \tag{2.54}$$

From this, the N-dependence of the bond dimension D_N required to achieve

$$\| \, |\psi\rangle - |\psi^D\rangle \, \| \le \frac{\epsilon_0}{N} \tag{2.55}$$

(as we demand that the error of physical quantities does not grow with N, hence the choice of $1/N$) is determined by

$$\frac{\epsilon_0}{N} \le (\text{const.}) \times D^{-\frac{1-\alpha}{\alpha}} N N^{\frac{c}{6}(1+\frac{1}{\alpha})}, \tag{2.56}$$

which leads to

$$D_N \sim \left(\frac{N^2}{\epsilon_0}\right)^{\frac{\alpha}{1-\alpha}} \times N^{\frac{c}{6}\frac{1+\alpha}{1-\alpha}}. \tag{2.57}$$

Thus, D_N is given as a polynomial function of the system size N.

3 Positive Maps and the Perron–Frobenius Theorem

In this section we summarize the necessary facts about positive maps. A compact and readable reference is [9]. The Perron–Frobenius type theorem for positive maps in finite dimensions can be found in [7].

3.1 Properties of Positive Maps

Definition 3.1. For $X \in \text{Mat}_n(\mathbb{C})$, we say X is positive semidefinite ($X \ge 0$) if $(v, Xv) \ge 0$ for all $v \in \mathbb{C}^n$. We say X is positive definite ($X > 0$) if $(v, Xv) > 0$ for all nonzero $v \in \mathbb{C}^n$.

In particular, if X is positive semidefinite, then X is Hermitian:

Proposition 3.2. *For $X \in \text{Mat}_n(\mathbb{C})$, if $(v, Xv) \in \mathbb{R}$ for all $v \in \mathbb{C}^n$, then X is Hermitian.*

(Proof) Note the identity (a type of polarization identity)

$$(x, Ay) = \frac{1}{4}\sum_{k=0}^{3} i^{-k}(x + i^k y, A(x + i^k y)). \tag{3.1}$$

Since $(v, Xv) = (X^\dagger v, v)$ and the assumption is that this is real, we have $(X^\dagger v, v) = (X^\dagger v, v)^* = (v, X^\dagger v)$. Thus, $(v, (X - X^\dagger)v) = 0$ for all $v \in \mathbb{C}^n$. By the above identity, it follows that $(x, (X - X^\dagger)y) = 0$ for all $x, y \in \mathbb{C}^n$. □

Hereafter, we always assume maps of the form

$$T : \text{Mat}_m(\mathbb{C}) \to \text{Mat}_n(\mathbb{C}) \tag{3.2}$$

are linear, i.e.

$$T(\alpha X + \beta Y) = \alpha T(X) + \beta T(Y). \tag{3.3}$$

Proposition 3.3. *If $X \geq 0$, then for any matrix A, $A^\dagger X A \geq 0$.*

(Proof)

$$\langle v|A^\dagger X A|v\rangle = (Av, XAv) \geq 0. \quad \square \tag{3.4}$$

Definition 3.4. A map $T : \mathrm{Mat}_m(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ is called positive if it preserves positive semidefiniteness:

$$X \geq 0 \;\Rightarrow\; T(X) \geq 0. \tag{3.5}$$

For example, given a family of matrices $\{A_i\}_{i=1}^d$ with $A_i \in \mathrm{Mat}_{m,n}(\mathbb{C})$,

$$T(X) = \sum_{i=1}^d A_i^\dagger X A_i \tag{3.6}$$

is positive.

Definition 3.5. For $T : \mathrm{Mat}_m(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$, the dual map $T^* : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_m(\mathbb{C})$ is defined by

$$\mathrm{tr}\,(Y T(X)) = \mathrm{tr}\,(T^*(Y) X), \quad X \in \mathrm{Mat}_m(\mathbb{C}),\; Y \in \mathrm{Mat}_n(\mathbb{C}). \tag{3.7}$$

Note that the dual map T^* is not the Hermitian adjoint $T^\dagger$ with respect to the Hilbert–Schmidt inner product $(X, Y)_{\mathrm{HS}} = \mathrm{tr}\,[X^\dagger Y]$. When X, Y are restricted to Hermitian matrices, however, T^* and $T^\dagger$ agree.

Proposition 3.6.

$$T \text{ is positive} \quad \Leftrightarrow \quad T^* \text{ is positive.}$$

(Proof) Assume T is positive. For the spectral decomposition $X = \sum_i \lambda_i |i\rangle\langle i|$,

$$\langle v|T^*(X)|v\rangle = \mathrm{tr}\,[T^*(X)\,|v\rangle\langle v|] = \mathrm{tr}\,[X\,T(|v\rangle\langle v|)] \tag{3.8}$$

$$= \sum_i \lambda_i \langle i|T(|v\rangle\langle v|)|i\rangle \geq 0, \tag{3.9}$$

hence T^* is positive. The converse is similar. $\square$

Definition 3.7. We say $T : \mathrm{Mat}_m(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ is trace-preserving if

$$\mathrm{tr}\,[T(X)] = \mathrm{tr}\,[X], \tag{3.10}$$

and unital if

$$T(1) = 1. \tag{3.11}$$

Proposition 3.8.

$$T \text{ is trace-preserving} \quad \Leftrightarrow \quad T^* \text{ is unital.}$$

(Proof) This follows from $\mathrm{tr}\,[T^*(1)X] = \mathrm{tr}\,[T(X)]$. □

Any Hermitian matrix X can be written as the difference of two positive semidefinite matrices:

Definition 3.9. For a Hermitian matrix $X \in \mathrm{Mat}_n(\mathbb{C})$, $X^\dagger = X$, with spectral decomposition $X = \sum_i \lambda_i |i\rangle\langle i|$, define

$$X_+ := \sum_{i,\lambda_i>0} \lambda_i |i\rangle\langle i|, \quad X_- := \sum_{i,\lambda_i<0} (-\lambda_i) |i\rangle\langle i|. \tag{3.12}$$

In particular,

$$X = X_+ - X_-, \quad |X| = \sqrt{X^\dagger X} = X_+ + X_-. \tag{3.13}$$

Proposition 3.10. *Any matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ can be expressed as a linear combination of positive semidefinite matrices.*

(Proof) Writing $A = (A + A^\dagger)/2 + i(A - A^\dagger)/(2i) = X + iY$, we obtain

$$A = X_+ - X_- + iY_+ - iY_-. \quad \square \tag{3.14}$$

Proposition 3.11. *Let $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ preserve Hermiticity, i.e.*

$$X^\dagger = X \ \Rightarrow \ T(X)^\dagger = T(X). \tag{3.15}$$

Then:

(i) For $A \in \mathrm{Mat}_n(\mathbb{C})$, we have $T(A)^\dagger = T(A^\dagger)$.

(ii) The eigenvalues of T are either real or occur in complex conjugate pairs (λ, λ^*).

(iii) For real eigenvalues, the eigenvectors of T can be chosen Hermitian.

(Proof) (i) Decompose $A = X + iY$ with $X^\dagger = X$, $Y^\dagger = Y$. Then

$$T(A)^\dagger = T(X)^\dagger - iT(Y)^\dagger = T(X) - iT(Y) = T(A^\dagger). \tag{3.16}$$

(ii) If $T(A) = \lambda A$, then $T(A^\dagger) = T(A)^\dagger = \lambda^* A^\dagger$. (iii) From $T(A + A^\dagger) = \lambda(A + A^\dagger)$. □

Proposition 3.12. *If $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ is positive, then T preserves Hermiticity.*

(Proof) Write $X = X_+ - X_-$ as the difference of positive semidefinite matrices. Then $T(X) = T(X_+) - T(X_-)$, but $T(X_+), T(X_-)$ are positive semidefinite and hence Hermitian. □

Therefore, the eigenvalues of a positive map T are either real or appear as complex conjugate pairs.

Next, we introduce matrix and operator norms.

Definition 3.13. For $v \in \mathbb{C}^n$, $A \in \mathrm{Mat}_n(\mathbb{C})$, and $T : \mathrm{Mat}_m(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$, define

$$\|v\| := \sqrt{\sum_{i=1}^{n} |v_i|^2}, \tag{3.17}$$

$$\|A\| := \sup_{v \in \mathbb{C}^n, \|v\|=1} \|Av\|, \tag{3.18}$$

$$\|T\| := \sup_{A \in \mathrm{Mat}_m(\mathbb{C}), \|A\|=1} \|T(A)\|. \tag{3.19}$$

Proposition 3.14. *The norm of a matrix A is given by its largest singular value:*

$$\|A\| = \sigma_{\max}(A). \tag{3.20}$$

(Proof) From the singular value decomposition $A = U\Sigma V^\dagger$, we have $\|Av\| = \|\Sigma v'\|$ with $v' = V^\dagger v$. □

Proposition 3.15.

$$\|Av\| \leq \|A\|\|v\|, \tag{3.21}$$

$$\|T(A)\| \leq \|T\|\|A\|. \tag{3.22}$$

(Proof)

$$\|A\| = \sup_{v \in \mathbb{C}^n, v \neq 0} \frac{\|Av\|}{\|v\|} \geq \frac{\|Av\|}{\|v\|}, \tag{3.23}$$

and similarly for $\|T\|$. □

Proposition 3.16.

$$\|AB\| \leq \|A\|\|B\|, \tag{3.24}$$

$$\|T_1 T_2\| \leq \|T_1\|\|T_2\|. \tag{3.25}$$

(Proof) For any $v \in \mathbb{C}^n$,

$$\|ABv\| \leq \|A\|\|Bv\| \leq \|A\|\|B\|\|v\|, \tag{3.26}$$

and similarly for $\|T_1 T_2\|$. □

Lemma 3.17.

$$\|A\| \leq 1 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & A \\ A^\dagger & 1 \end{pmatrix} \geq 0. \tag{3.27}$$

(Proof) Let $A = U\Sigma V^\dagger$ be the singular value decomposition with $\sigma_1 \geq \cdots \geq \sigma_n$. Then

$$\begin{pmatrix} 1 & A \\ A^\dagger & 1 \end{pmatrix} = \begin{pmatrix} U & \\ & V \end{pmatrix} \bigoplus_i \begin{pmatrix} 1 & \sigma_i \\ \sigma_i & 1 \end{pmatrix} \begin{pmatrix} U^\dagger & \\ & V^\dagger \end{pmatrix}. \tag{3.28}$$

Hence the eigenvalues are $\{1+\sigma_i, 1-\sigma_i\}_i$. Therefore,

$$\begin{pmatrix} 1 & A \\ A^\dagger & 1 \end{pmatrix} \geq 0 \quad \Leftrightarrow \quad \sigma_1 \leq 1 \quad \Leftrightarrow \quad \|A\| \leq 1. \quad \square \tag{3.29}$$

Theorem 3.18 (Russo–Dye)**.** *Let $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ be positive and unital. Then*

$$\|T\| = 1. \tag{3.30}$$

(Proof) For a general unitary matrix $U \in U(n)$, take its spectral decomposition

$$U = \sum_i e^{i\phi_i} P_i, \quad P_i^\dagger = P_i, \quad P_i^2 = P_i. \tag{3.31}$$

Then

$$\begin{pmatrix} 1 & T(U) \\ T(U)^\dagger & 1 \end{pmatrix} = \sum_i \begin{pmatrix} T(P_i) & e^{i\phi_i} T(P_i) \\ e^{-i\phi_i} T(P_i) & T(P_i) \end{pmatrix} = \sum_i \begin{pmatrix} 1 & e^{i\phi_i} \\ e^{-i\phi_i} & 1 \end{pmatrix} \otimes T(P_i). \tag{3.32}$$

Each term in the last expression is positive semidefinite[8]. Thus

$$\begin{pmatrix} 1 & T(U) \\ T(U)^\dagger & 1 \end{pmatrix} \geq 0. \tag{3.33}$$

By the preceding lemma, this implies $\|T(U)\| \leq 1$.

Now, for a general $A \in \mathrm{Mat}_n(\mathbb{C})$ with $\|A\| = 1$, take its singular value decomposition $A = U\Sigma V^\dagger$ with $1 = \sigma_1 \geq \sigma_2 \geq \cdots$. Since

$$\sigma_i = \cos\theta_i = \frac{e^{i\theta_i} + e^{-i\theta_i}}{2}, \tag{3.34}$$

we can write

$$A = \frac{1}{2} U \,\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n}) V^\dagger + \frac{1}{2} U \,\mathrm{diag}(e^{-i\theta_1}, \ldots, e^{-i\theta_n}) V^\dagger =: \frac{1}{2} U_1 + \frac{1}{2} U_2, \tag{3.35}$$

as a convex combination of two unitary matrices. Therefore

$$\|T(A)\| \leq \tfrac{1}{2}(\|T(U_1)\| + \|T(U_2)\|) \leq 1. \tag{3.36}$$

On the other hand,

$$\|T\| \geq \|T(1)\| = \|1\| = 1. \tag{3.37}$$

Hence $\|T\| = 1$. $\square$

Corollary 3.19 (Russo–Dye)**.** *For a positive map $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$,*

$$\|T\| = \|T(1)\|. \tag{3.38}$$

(Proof) Let $P := T(1) \geq 0$. If $P > 0$, then the map $T'(A) = P^{-1/2} T(A) P^{-1/2}$ is unital with $T'(1) = 1$. Thus,

$$\|T(A)\| = \|P^{1/2} T'(A) P^{1/2}\| \leq \|P\| \|T'(A)\| \leq \|P\| \|T'\| \|A\| = \|P\| \|A\|. \tag{3.39}$$

[8] If $A, B \geq 0$, then the eigenvalues of $A \otimes B$ are $\lambda_i(A)\lambda_j(B) \geq 0$, hence $A \otimes B \geq 0$.

Hence $\|T\| \leq \|P\|$. On the other hand, $\|T(1)\| = \|P\|$, so $\|T\| = \|P\| = \|T(1)\|$.

If P has zero eigenvalues, then for small $\epsilon > 0$, define $T_\epsilon(A) = T(A) + \epsilon 1$. Then $T_\epsilon(1) = P + \epsilon 1 > 0$. Thus $\|T_\epsilon\| = \|T_\epsilon(1)\|$. Taking $\epsilon \to 0$, we obtain the desired result. □

Definition 3.20 (Trace Norm)**.** For $A \in \mathrm{Mat}_n(\mathbb{C})$, the trace norm is defined by

$$\|A\|_{\mathrm{tr}} = \mathrm{tr}\,|A|. \tag{3.40}$$

If $\{\sigma_a\}_a$ are the singular values of A, then

$$\|A\|_{\mathrm{tr}} = \sum_{a=1}^{n} \sigma_a. \tag{3.41}$$

Note also that for any $U, V \in U(n)$,

$$\|UAV^\dagger\|_{\mathrm{tr}} = \|A\|_{\mathrm{tr}}. \tag{3.42}$$

Next, we prove the contractivity of the trace norm under positive maps. As preparation, we first prove the following lemma.

Lemma 3.21. *For $A, B \in \mathrm{Mat}_n(\mathbb{C})$,*

$$|\mathrm{tr}\,(AB)| \leq \|A\|_{\mathrm{tr}}\|B\|. \tag{3.43}$$

Equality holds when $A = U|A|$ is the polar decomposition and $B = U^\dagger$.

(Proof) Take the polar decomposition $A = U|A|$[9]. Define the linear map

$$\phi_{|A|} : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C}), \quad \phi_{|A|}(B) := \mathrm{tr}\,(|A|B) \cdot 1_n = \mathrm{tr}\,(\sqrt{|A|}B\sqrt{|A|}) \cdot 1_n. \tag{3.44}$$

This map is positive, so by Russo–Dye, for any $B \in \mathrm{Mat}_n(\mathbb{C})$,

$$|\mathrm{tr}\,(|A|B)| = \|\phi_{|A|}(B)\| \leq \|\phi_{|A|}\| \cdot \|B\| \leq \|\phi_{|A|}(1)\| \cdot \|B\| = \mathrm{tr}\,|A| \cdot \|B\| = \|A\|_{\mathrm{tr}} \cdot \|B\|. \tag{3.45}$$

Therefore,

$$|\mathrm{tr}\,(AB)| = |\mathrm{tr}\,(|A|BU)| \leq \|A\|_{\mathrm{tr}} \cdot \|BU\| = \|A\|_{\mathrm{tr}} \cdot \|B\|. \tag{3.46}$$

Equality holds when $BU = 1_n$. □

From this, we obtain the variational expression for the trace norm:

Corollary 3.22.

$$\|A\|_{\mathrm{tr}} = \sup_{B, \|B\|=1} |\mathrm{tr}\,(AB)|. \tag{3.47}$$

Now we can state the following.

[9] If $A = U\Sigma V^\dagger$ is the singular value decomposition, then $A = (UV^\dagger)V\Sigma V^\dagger$ gives the polar decomposition.

Theorem 3.23 (Contractivity of the Trace Norm)**.** *Let* $T : \mathrm{Mat}_m(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ *be positive. Then for any* $A \in \mathrm{Mat}_m(\mathbb{C})$*,*

$$\|T(A)\|_{\mathrm{tr}} \leq \|T^*(1)\| \cdot \|A\|_{\mathrm{tr}}. \tag{3.48}$$

In particular, if T *is trace-preserving,*

$$\|T(A)\|_{\mathrm{tr}} \leq \|A\|_{\mathrm{tr}}. \tag{3.49}$$

(Proof)

$$\|T(A)\|_{\mathrm{tr}} = \sup_{B, \|B\|=1} |\mathrm{tr}\,(T(A)B)| = \sup_{B, \|B\|=1} |\mathrm{tr}\,(AT^*(B))|. \tag{3.50}$$

Here,

$$|\mathrm{tr}\,(AT^*(B))| \leq \|A\|_{\mathrm{tr}} \cdot \|T^*(B)\| \leq \|A\|_{\mathrm{tr}} \cdot \|T^*\| \cdot \|B\| \tag{3.51}$$
$$= \|A\|_{\mathrm{tr}} \cdot \|T^*(1)\| \cdot \|B\|. \tag{3.52}$$

Thus

$$\|T(A)\|_{\mathrm{tr}} \leq \|A\|_{\mathrm{tr}} \cdot \|T^*(1)\|. \tag{3.53}$$

□

3.2 Perron–Frobenius Theorem

We prove the Perron–Frobenius theorem for finite-dimensional positive linear maps, following [7]. This is the finite-dimensional version of the Krein–Rutman theorem, an extension of the Perron–Frobenius theorem to Banach spaces.

Definition 3.24. For $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$, define

- $\mathrm{Sp}(T) := \{\lambda \in \mathbb{C} \mid \lambda - T \text{ is not invertible}\}$, the spectrum of T.
- $\rho_T = \max_{\lambda \in \mathrm{Sp}(T)} |\lambda|$, the spectral radius of T.

The statements we wish to prove are as follows. For a positive linear map $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$:

- $\rho_T \in \mathrm{Sp}(T)$.
- There exists $^\exists X \geq 0$ such that $T(X) = \rho_T X$.

That is, for a positive map the spectral radius is an eigenvalue, and moreover an eigenvector can be chosen to be positive.

Because the proof is somewhat long, we first outline the steps:

- When T satisfies the good condition of being "irreducible," we prove a stronger statement.
- We introduce the notion of *hereditary*.
- If T is irreducible, then $(1 + T)^{n-1} > 0$.
- If T is irreducible, there exists a unique (up to scale) eigenpair $T(X) = \rho_T X$ with $X > 0$.

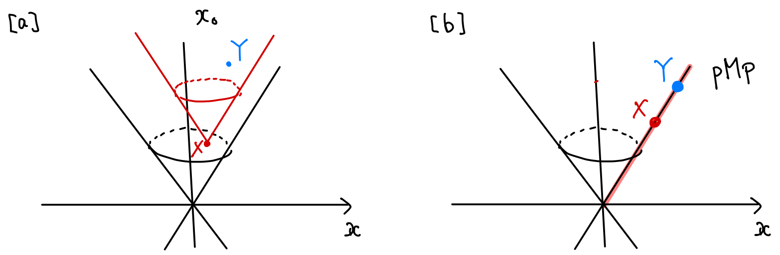

Figure 2

- A general positive map is obtained as a limit of irreducible positive maps.

In this section, unless a specific n must be indicated, we abbreviate $M = \mathrm{Mat}_n(\mathbb{C})$, and linear maps $T : M \to M$ are assumed positive. We also denote by

$$M_+ := \{X \in M \mid X \geq 0\} \tag{3.54}$$

the cone of positive semidefinite matrices.

Definition 3.25. A subalgebra $M' \subset M$ is called *hereditary* if it has the following property:

$$0 \leq X \leq Y \text{ and } Y \in M' \;\Rightarrow\; X \in M'. \tag{3.55}$$

Proposition 3.26. *Let $p \in M$ be an orthogonal projection. Then the subalgebra pMp is hereditary.*

Before the proof, let us look at an example. Let $M = \mathrm{Mat}_2(\mathbb{C})$. Then the cone of positive semidefinite matrices is

$$M_+ = \{x_0 + \boldsymbol{x} \cdot \boldsymbol{\sigma} \mid x_0 \geq |\boldsymbol{x}|\}, \tag{3.56}$$

namely the interior and boundary of a cone. The condition $0 \leq X \leq Y$ means that X lies in the cone with apex at the origin, and Y lies in the cone with apex at X (see Fig. 2 [a]). Consider the orthogonal projection onto the $(1,1)$-component, $p = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then

$$(pMp)_+ = \{x_0 + x_3\sigma_3 \mid x_0 = x_3\}, \tag{3.57}$$

so $Y \in (pMp)_+$ sits on the "rim" of the cone M_+. For $Y \geq X$ to hold, X must also lie on the rim of M_+, which shows $X \in (pMp)_+$ (see Fig. 2 [b]).

(Proof. [10]) Assume $0 \leq X \leq pYp$ with $Y \in M$. Then

$$0 \leq (1-p)X(1-p) \leq (1-p)pYp(1-p) = 0, \tag{3.58}$$

hence $(1-p)\sqrt{X} = \sqrt{X}(1-p) = 0$. Thus $(1-p)X = X(1-p) = 0$, and therefore $X \in pMp$. □

We define irreducibility of a positive linear map T as follows.

Definition 3.27. A map T is *irreducible* if there exists no nontrivial (i.e., neither $p = 0$ nor $p = 1$) orthogonal projection p such that

$$T(pMp) \subset pMp. \tag{3.59}$$

Recall that an irreducible representation $\rho : G \to V$ of a group G (with representation space V) is one for which ρ_g cannot be put in block-diagonal form, i.e., there is no nontrivial projection p with $\rho_g(pV) \subset pV$. The above generalizes the notion of irreducibility in the Perron–Frobenius theorem to positive linear maps.

The condition $T(pMp) \subset pMp$ can be reformulated as a property of p:

Lemma 3.28. *Let $p \in M$ be an orthogonal projection. Then*

$$T(pMp) \subset pMp \quad \Leftrightarrow \quad {}^{\exists}\lambda > 0 \ s.t.\ T(p) \le \lambda p. \tag{3.60}$$

(Proof[10]) Choose a basis so that $p = \begin{pmatrix} 1_r & \\ & 0 \end{pmatrix}$. Let $X \in pMp$, i.e.,

$$X = \begin{pmatrix} X' & 0 \\ 0 & 0 \end{pmatrix}, \quad X' \in \mathrm{Mat}_r(\mathbb{C}). \tag{3.61}$$

By linearity of T and the fact that any matrix is a linear combination of positive semidefinite matrices, we may assume $X \ge 0$. Write the spectral decomposition

$$X = \sum_{a=1}^{q} \lambda_a p_a, \quad \lambda_1 \ge \cdots \ge \lambda_q > 0, \quad q \le r, \tag{3.62}$$

and note that

$$p = \sum_{a=1}^{r} p_a. \tag{3.63}$$

Then

$$0 \le \lambda_q p \le X \le \lambda_1 p. \tag{3.64}$$

($\Leftarrow$) For any $X \ge 0$, $X \ne 0$, there exists λ_1 such that $0 \le T(X) \le \lambda_1 T(p) \le \lambda_1 \lambda p \in (pMp)_+$. Hence, by heredity, $T(X) \in pMp$.

($\Rightarrow$) For some $X \ge 0$, $X \ne 0$, we have $0 \le \lambda_q T(p) \le T(X) \in (pMp)_+$. Thus, by heredity, $\lambda_q T(p) \in (pMp)_+$, i.e., $T(p) \in (pMp)_+$. Since $T(p) \ge 0$ and $T(p) \in pMp$, its spectral decomposition

$$T(p) = \sum_{a=1}^{s} \xi_a p'_a, \quad \xi_1 \ge \cdots \ge \xi_s > 0, \quad s \le r, \tag{3.65}$$

implies $T(p) \le \xi_1 p$. □

The following statements hold.

Proposition 3.29. *Assume T is irreducible. Then*

(i) $T(1) > 0$.

(ii) If $X > 0$, then $T(X) > 0$.

(Proof) (i) Suppose $T(1)$ has a zero eigenvalue. Let p be the orthogonal projection onto $\mathrm{im}\, T(1)$. Since $T(p) \le T(1) \in pMp$, there exists $\lambda > 0$ with $T(1) \le \lambda p$. This contradicts the irreducibility of T.

(ii) If $X > 0$, there exists $\epsilon > 0$ such that $X > \epsilon 1_n$. Hence $T(X) > \epsilon T(1) > 0$. □

[10] I learned this from Kan Kitamura; many thanks.

We next reformulate the irreducibility of T.

Definition 3.30. We say that T is *strictly positive* (strictly irreducible), written $T > 0$, if for every $X \geq 0$, $X \neq 0$, one has $T(X) > 0$.

Define the Hilbert–Schmidt inner product by

$$(A, B)_{\mathrm{HS}} := \mathrm{tr}\,[A^\dagger B], \quad A, B \in \mathrm{Mat}_n(\mathbb{C}). \tag{3.66}$$

Proposition 3.31.

- *If* $X, Y \geq 0$*, then* $(X, Y)_{\mathrm{HS}} \geq 0$.
- *If* $X > 0$ *and* $Y \geq 0$ *(or* $X \geq 0$ *and* $Y > 0$*), then* $(X, Y)_{\mathrm{HS}} > 0$.
- $T > 0$ *is equivalent to* $(Y, T(X))_{\mathrm{HS}} > 0$ *for all* $X, Y \geq 0$ *with* $X, Y \neq 0$.

(Proof) The first two items follow from $\mathrm{tr}\,[XY] = \mathrm{tr}\,[\sqrt{Y} X \sqrt{Y}] \geq 0$. For the third: ($\Rightarrow$) since $\mathrm{tr}\,[Y T(X)] = \mathrm{tr}\,[\sqrt{T(X)} Y \sqrt{T(X)}]$ and $T > 0$ with $X \neq 0$ implies $T(X) > 0$ (positive definite), we have $\mathrm{tr}\,[\sqrt{T(X)} Y \sqrt{T(X)}] > 0$ for $Y \neq 0$. ($\Leftarrow$) For any $v \in \mathbb{C}^n$, set $Y = |v\rangle\langle v|$. Then $\langle v|T(X)|v\rangle > 0$, hence $T(X) > 0$. □

Lemma 3.32.

$$T \text{ is irreducible} \quad \Leftrightarrow \quad (1 + T)^{n-1} > 0. \tag{3.67}$$

(Proof) ($\Rightarrow$) For $Y \geq 0$, $Y \neq 0$, consider $(1 + T)^{n-1}(Y)$. If $Y > 0$ then $(1 + T)(Y) > 0$, so assume Y has zero eigenvalues. If $(1 + T)(Y)v = 0$ for some $v \in \mathbb{C}^n$, then $\langle v|(1 + T)(Y)|v\rangle = 0$ implies $\langle v|Y|v\rangle = 0$, hence $Yv = 0$. Thus $\ker(1 + T)(Y) \subset \ker Y$ always holds.

Suppose $\ker(1+T)(Y) = \ker Y$, i.e., also $\ker Y \subset \ker(1+T)(Y)$. If $Yv = 0$, then $(1+T)(Y)v = T(Y)v$, so $\ker Y \subset \ker T(Y)$, equivalently $\mathrm{im}\, T(Y) \subset \mathrm{im}\, Y$. Let p be the orthogonal projection onto $\mathrm{im}\, Y$ (if $Y = \sum_{a=1}^r \lambda_a p_a$ with $r < n$, then $p = \sum_{a=1}^r p_a$). Since $T(p) \in \mathrm{im}\, Y = \mathrm{im}\, p$, letting ξ_1 be the largest eigenvalue of $T(p)$, we have $T(p) \leq \xi_1 p$. By Lemma 3.28, $T(pMp) \subset pMp$. As T is irreducible and $Y \neq 0$, this forces $p = 1$, a contradiction because then Y would be invertible. Hence $\ker(1 + T)(Y) \neq \ker Y$, so $\ker(1 + T)(Y) \subsetneq \ker Y$, i.e., $\mathrm{rank}(1 + T)(Y) > \mathrm{rank} Y$. Applying $(1 + T)$ at least $n - 1$ times to a singular Y yields $(1 + T)^{n-1}(Y)$ full rank, i.e., $(1 + T)^{n-1}(Y) > 0$.

($\Leftarrow$) By contradiction. If T is reducible, there exists a nontrivial projection p and $\lambda > 0$ with $T(p) \leq \lambda p$. Then

$$0 < (1 + T)^{n-1}(p) \leq (1 + \lambda)^{n-1} p, \tag{3.68}$$

whose right-hand side is not full rank—a contradiction. □

As an example of an irreducible positive map:

- Example. $T(A) := \frac{1}{n}\mathrm{tr}\,[A]\, 1_n$ is irreducible, since $T^2 = T$ and hence $(1+T)^{n-1} = 1+(\mathrm{const.})\, T > 0$.

We introduce yet another equivalent formulation of irreducibility.

Lemma 3.33. *T is irreducible if and only if the following holds:*

- *For any nonzero positive semidefinite $X, Y \geq 0$ with $(Y, X)_{\mathrm{HS}} = 0$, there exists $k \in \mathbb{N}$ such that $(Y, T^k(X))_{\mathrm{HS}} > 0$.*

This corresponds to the graph-theoretic condition in the classical Perron–Frobenius theorem: for any pair of nodes (i, j) there exists a path $ii_1 \to i_1 i_2 \to \cdots \to i_k j$ through nonzero matrix entries.

(Proof) ($\Rightarrow$) From irreducibility, for $X, Y \geq 0$, $X, Y \neq 0$, we have $(Y, (1+T)^{n-1}(X))_{\mathrm{HS}} > 0$. Expanding and using $(Y, X)_{\mathrm{HS}} = 0$,

$$\left(Y, \left\{(n-1)T + \frac{1}{2}(n-1)(n-2)T^2 + \cdots + T^{n-1}\right\}(X)\right)_{\mathrm{HS}} > 0. \tag{3.69}$$

All terms are nonnegative, so there exists at least one $k \in \{1, \ldots, n-1\}$ with $(Y, T^k(X))_{\mathrm{HS}} > 0$.

($\Leftarrow$) Prove the contrapositive. Suppose there exists a nontrivial projection p with $T(pMp) \subset pMp$. Then for any $k > 0$, $T^k(p) \in pMp$, hence $T^k(p)\, p = T^k(p)$. Therefore, for all $k > 0$,

$$0 = \mathrm{tr}\left[(1-p)T^k(p)\right] = (1-p, T^k(p))_{\mathrm{HS}}. \tag{3.70}$$

Taking $X = p$, $Y = 1-p$—a nonzero pair with $(Y, X)_{\mathrm{HS}} = 0$—we have $(Y, T^k(X))_{\mathrm{HS}} = 0$ for all $k \in \mathbb{N}$. ☐

Since for Hermitian matrices X, Y the dual satisfies $(X, T^k(Y))_{\mathrm{HS}} = (Y, T^{*k}(X))_{\mathrm{HS}}$, we obtain:

Corollary 3.34.

$$T \text{ is irreducible} \quad \Leftrightarrow \quad T^* \text{ is irreducible.} \tag{3.71}$$

Finally, define the following real-valued function on M_+:

$$r : M_+ \to \mathbb{R}, \quad r(X) := \sup\{\lambda \in \mathbb{R} \mid T(X) \geq \lambda X\}. \tag{3.72}$$

Define also the overall supremum

$$r := \sup_{X \in M_+} r(X). \tag{3.73}$$

The Perron–Frobenius theorem follows from the next lemma.

Lemma 3.35 ([7])**.** *Assume T is irreducible. Then*

(i) There exists $X = Z$ attaining the supremum r and Z is positive definite, $Z > 0$.

(ii) $T(Z) = rZ$.

(iii) The eigenspace for $\lambda = r$ is one-dimensional.

(iv) $r = \rho_T$ (the spectral radius).

(Proof) (i) We first show the existence of $X \in M_+$ attaining the supremum. Since $r(aX) = a\, r(X)$ for $a > 0$, we may assume $\mathrm{tr}\,[X] = 1$. The "sphere"

$$S = \{X \in M_+ \mid \mathrm{tr}\,[X] = 1\} \tag{3.74}$$

is compact. Thus, if the restriction $r|_S$ were continuous, it would attain a maximum on S. However, r need not be continuous on S.

On the other hand, for $X > 0$ we have $T(X) > 0$ (by Proposition 3.29), and then

$$r(X) = \left\| T(X)^{-\frac{1}{2}} X \, T(X)^{-\frac{1}{2}} \right\|^{-1}, \tag{3.75}$$

so r is continuous on the set of positive definite matrices $\{X \in M_+ \mid X > 0\}$. To see that the supremum r is achieved by some $X > 0$, introduce

$$N := (1+T)^{n-1} S = \{(1+T)^{n-1}(X) > 0 \mid \mathrm{tr}\,[X] = 1\}. \tag{3.76}$$

The set N is compact and consists only of positive definite matrices. Hence $r|_N$ attains its maximum at some $Z > 0$. We claim that Z attains the supremum over all of M_+.

Take $X \in S$ and set $Y = (1+T)^{n-1}(X) \in N$. By definition of $r(X)$, $T(X) - r(X)X \geq 0$. Hence

$$\begin{aligned} (T - r(X))(Y) &= (T - r(X))(1+T)^{n-1}(X) && (3.77)\\ &= (1+T)^{n-1}\big(T(X) - r(X)X\big) \geq 0. && (3.78) \end{aligned}$$

Thus $T(Y) \geq r(X)\, Y$, and therefore

$$r(Y) \geq r(X) \quad \text{for all } X \in S. \tag{3.79}$$

Consequently,

$$r \geq \max_{X \in N} r(X) \geq \sup_{X \in S} r(X) = \sup_{X \in M_+} r(X) = r, \tag{3.80}$$

whence

$$r = \max_{X \in N} r(X). \tag{3.81}$$

Therefore the supremum r is attained at some positive definite matrix $Z > 0$.

(ii) Suppose, to the contrary, that $T(Z) - rZ \neq 0$. Let $W = (1+T)^{n-1}(Z)$. Then

$$T(W) - rW = (1+T)^{n-1}\big(T(Z) - rZ\big) > 0, \tag{3.82}$$

which implies $r(W) > r$, contradicting the maximality of r.

(iii) Suppose there exists another eigenvector Z' not proportional to Z with $T(Z') = rZ'$. Since r is real, Z' can be taken Hermitian (Proposition 3.11). Consider the spectral decomposition

$$Z^{-\frac{1}{2}} Z' Z^{-\frac{1}{2}} = \sum_{a=1}^{n} \lambda_a p_a, \quad \lambda_1 \geq \cdots \geq \lambda_n. \tag{3.83}$$

Then

$$\lambda_1 1 - Z^{-\frac{1}{2}} Z' Z^{-\frac{1}{2}} = \sum_{a=2}^{n} (\lambda_1 - \lambda_a) p_a \geq 0 \tag{3.84}$$

is not full rank. Hence $\lambda_1 Z - Z'$ is not full rank; but

$$0 < (1+T)^{n-1}(\lambda_1 Z - Z') = (1+r)^{n-1}(\lambda_1 Z - Z'), \tag{3.85}$$

a contradiction.

(iv) Define

$$\tilde{T}(A) := \frac{1}{r} Z^{-\frac{1}{2}} \, T\Big(Z^{\frac{1}{2}} A Z^{\frac{1}{2}}\Big) \, Z^{-\frac{1}{2}}. \tag{3.86}$$

Then $\tilde{T}(1) = 1$, so $\tilde{T}$ is unital. By Theorem 3.18, $\|\tilde{T}\| = 1$. If $T(A) = \alpha A$, then

$$\tilde{T}\left(Z^{-\frac{1}{2}} A Z^{-\frac{1}{2}}\right) = \frac{\alpha}{r} Z^{-\frac{1}{2}} A Z^{-\frac{1}{2}}. \tag{3.87}$$

Therefore

$$1 = \|\tilde{T}\| \geq \left|\frac{\alpha}{r}\right|. \tag{3.88}$$

Thus $|\alpha| \leq r$ for every eigenvalue α, i.e. $r = \rho_T$. □

Moreover, any eigenvector corresponding to an eigenvalue $\lambda \neq r$ cannot be positive semidefinite.

Lemma 3.36 ([7]). *Assume T is irreducible. If $Y \geq 0$, $Y \neq 0$ is an eigenvector of T with eigenvalue α, then $\alpha = r$.*

(Proof) Use the dual T^*. Then T^* is also irreducible. As for T, define

$$r^* := \sup_{X \in M_+} \sup\{\lambda \in \mathbb{R} \mid T^*(X) \geq \lambda X\}. \tag{3.89}$$

There exists $Z^* > 0$ such that $T^*(Z^*) = r^* Z^*$. Then

$$r\,(Z^*, Z)_{\mathrm{HS}} = (Z^*, T(Z))_{\mathrm{HS}} = (T^*(Z^*), Z)_{\mathrm{HS}} = r^*(Z^*, Z)_{\mathrm{HS}} > 0, \tag{3.90}$$

hence $r = r^*$. If $T(Y) = \alpha Y$ with $Y \geq 0$, $Y \neq 0$, then

$$\alpha\,(Z^*, Y)_{\mathrm{HS}} = (Z^*, T(Y))_{\mathrm{HS}} = (T^*(Z^*), Y)_{\mathrm{HS}} = r\,(Z^*, Y)_{\mathrm{HS}}. \tag{3.91}$$

Since $Z^* > 0$, we have $(Z^*, Y)_{\mathrm{HS}} > 0$ (Proposition 3.31), hence $\alpha = r$. □

Putting everything together, we obtain:

Theorem 3.37 ([7]. (Finite-dimensional Krein–Rutman Theorem)). *Let $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ be an irreducible positive linear map. Then*

(i) $\rho_T \in \mathrm{Sp}(T)$.

(ii) The eigenspace for $\lambda = \rho_T$ is nondegenerate and has a positive definite eigenvector:

$$T(Z) = \rho_T Z, \quad Z > 0. \tag{3.92}$$

(iii) Any eigenvector corresponding to an eigenvalue $\lambda \neq \rho_T$ is not positive semidefinite.

For an irreducible positive linear map T, the spectrum has the shape illustrated in Fig. 3.

Corollary 3.38. *Let $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ be an irreducible positive linear map. Then*

(i) $\rho_T \in \mathrm{Sp}(T)$.

(ii) An eigenvector for the eigenvalue $\lambda = \rho_T$ can be chosen positive semidefinite.

(Proof [5]) The set of irreducible positive linear maps is dense in the set of all positive linear maps. Indeed, given any positive linear map T and some irreducible positive linear map S, the map

$$T_\epsilon = T + \epsilon S \tag{3.93}$$

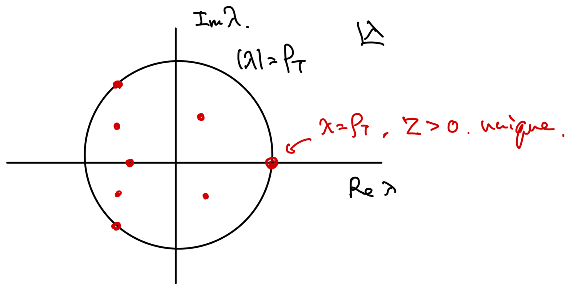

Figure 3

is irreducible because

$$(1+T_\epsilon)^{n-1} = (1+\epsilon S)^{n-1} + \cdots > 0. \tag{3.94}$$

(Recall that if T_1, T_2 are positive, then so is $T_1 \circ T_2$.) Thus, for any T there exist irreducible T_ϵ arbitrarily close to T. Applying Theorem 3.37 to irreducible T_ϵ, we have

$$T_\epsilon(Z_\epsilon) = \rho_{T_\epsilon} Z_\epsilon, \quad Z_\epsilon > 0. \tag{3.95}$$

Taking $\epsilon \to +0$,

$$T(Z) = \rho_T Z, \quad Z \geq 0. \tag{3.96}$$

(Here Z need not be positive definite, and the eigenspace for $\lambda = \rho_T$ need not be nondegenerate.) □

Irreducibility is preserved under the following "similarity transformations."

Proposition 3.39. *Let $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ be an irreducible positive linear map. For any $c > 0$ and any $Q \in \mathrm{GL}_n(\mathbb{C})$, define*

$$\tilde{T}(A) = c\, Q^{-1} T(QAQ^\dagger)(Q^\dagger)^{-1}. \tag{3.97}$$

Then

(i) $\tilde{T}$ is irreducible.

(ii) $\tilde{T}^*(A) = c\, Q^\dagger\, T^*\big((Q^{-1})^\dagger A Q^{-1}\big)\, Q$.

(Proof) (i) Let q be an orthogonal projection and assume $\tilde{T}(q) \leq \lambda q$ for some $\lambda > 0$. This is equivalent to $c\, T(QqQ^\dagger) \leq \lambda\, QqQ^\dagger$. Let p be the orthogonal projection onto $\mathrm{im}\,(QqQ^\dagger)$; then p "reduces" T. Indeed, with the spectral decomposition

$$QqQ^\dagger = \sum_{a=1}^{r} \lambda_a p_a, \quad \lambda_1 \geq \cdots \geq \lambda_r > 0, \tag{3.98}$$

and $p = \sum_{a=1}^{r} p_a$, we have $\lambda_r p \leq QqQ^\dagger \leq \lambda_1 p$ and hence $\lambda_r T(p) \leq \lambda_1 p$. Thus $c\, \lambda_r T(p) \leq \lambda\, \lambda_r p$. Since T is irreducible, $p = 0$ or $p = 1$.

(ii) A straightforward computation:

$$
\begin{aligned}
(\tilde{T}^*(A), B)_{\mathrm{HS}} &= (A, \tilde{T}(B))_{\mathrm{HS}} = \mathrm{tr}\big[A^\dagger\, c\, Q^{-1} T(QBQ^\dagger)(Q^\dagger)^{-1}\big] && (3.99)\\
&= \mathrm{tr}\big[c\,(Q^\dagger)^{-1} A^\dagger Q^{-1}\, T(QBQ^\dagger)\big] = \mathrm{tr}\big[c\, T^*\big((Q^\dagger)^{-1} A^\dagger Q^{-1}\big)\, QBQ^\dagger\big] && (3.100)\\
&= \mathrm{tr}\big[c\, Q^\dagger T^*\big((Q^\dagger)^{-1} A^\dagger Q^{-1}\big)\, QB\big], && (3.101)
\end{aligned}
$$

which yields the claimed formula. □

Proposition 3.40 ([2, 5]. Normalization)**.** *Let* $T : \mathrm{Mat}_n(\mathbb{C}) \to \mathrm{Mat}_n(\mathbb{C})$ *be an irreducible positive linear map. By a similarity transform one can achieve*

(i) $T(1) = 1$*,*

(ii) $T^*(\Lambda) = \Lambda$*, with* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$*,* $\lambda_1 \geq \cdots \geq \lambda_n > 0$*, and* $\mathrm{tr}\,[\Lambda] = 1$*.*

(Proof) Let

$$T(Z) = \rho_T Z > 0, \quad T^*(Z^*) = \rho_T Z^* > 0. \tag{3.102}$$

Consider the similarity transform

$$T_1(A) := \rho_T^{-1}\, Z^{-1/2}\, T(Z^{1/2} A Z^{1/2})\, Z^{-1/2}. \tag{3.103}$$

Then $T_1(1) = 1$, and

$$T_1^*\big(Z^{1/2} Z^* Z^{1/2}\big) = Z^{1/2} Z^* Z^{1/2}. \tag{3.104}$$

Diagonalize $Z^{1/2} Z^* Z^{1/2} = U \Lambda U^\dagger$, and set

$$T_2(A) := U^\dagger\, T_1(U A U^\dagger)\, U. \tag{3.105}$$

Then $T_2(1) = 1$ and $T_2^*(\Lambda) = \Lambda$. (Rescale to ensure $\mathrm{tr}\,[\Lambda] = 1$.) □

4 TI-MPS

From this section on, we study MPS with translational symmetry.

4.1 Existence and Canonical Form of TI-MPS

When a Hamiltonian for a 1D spin system is defined with periodic boundary conditions, its eigenstates are expected to possess translational symmetry in an appropriate sense. Hence one may expect that the site dependence of the matrices $\{A_i^{[m]}\}$ can be removed.

Definition 4.1. An MPS whose matrices $A_i^{[m]}$ do not depend on m is called a translationally invariant (TI) MPS.

Define the translation operator by

$$\hat{\mathrm{Tr}}\,|i_1 \cdots i_N\rangle := |i_N i_1 \cdots i_{N-1}\rangle\,. \tag{4.1}$$

In this note we consider only states with Bloch momentum 0, i.e., states $|\psi\rangle$ satisfying $\hat{\mathrm{Tr}}\,|\psi\rangle = |\psi\rangle$.

Theorem 4.2. *Any state with* $\hat{\mathrm{Tr}}\,|\psi\rangle = |\psi\rangle$ *admits a TI-MPS representation.*

Logically, this is a nontrivial statement.

(Proof) Let an OBC-MPS of $|\psi\rangle$ be

$$|\psi\rangle = A^{[1]}_{i_1} \cdots A^{[N]}_{i_N} |i_1 \cdots i_N\rangle . \tag{4.2}$$

By translation invariance, for any k,

$$|\psi\rangle = \operatorname{tr}[A^{[k+1]}_{i_1} \cdots A^{[k+N]}_{i_N}] |i_1 \cdots i_N\rangle . \tag{4.3}$$

Average over k. The following matrices yield a TI-MPS:

$$B_i := N^{-\frac{1}{N}} \begin{pmatrix} 0 & A^{[1]}_i & & & \\ & 0 & A^{[2]}_i & & \\ & & \ddots & & \\ & & & 0 & A^{[N-1]}_i \\ A^{[N]} & & & & 0 \end{pmatrix} . \quad \square \tag{4.4}$$

Thus, a translationally invariant state $\hat{\operatorname{Tr}}|\psi\rangle = |\psi\rangle$ can be represented by a set $\{A_i\}_{i=1}^d$ of square matrices at a single site:

$$|\psi\rangle = \sum_{i_1 \ldots i_N} \operatorname{tr}[A_{i_1} \cdots A_{i_N}] |i_1 \cdots i_N\rangle . \tag{4.5}$$

We would like to classify the matrices $\{A_i\}_i$ in order to classify "gapped, nondegenerate ground states." We also want to exclude the following kinds of states:

- Example.

  $$A_\uparrow = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \quad A_\downarrow = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}. \tag{4.6}$$

  The resulting state is a superposition of macroscopic states,

  $$|\psi\rangle = |\uparrow \cdots \uparrow\rangle + |\downarrow \cdots \downarrow\rangle . \tag{4.7}$$

  More generally, we wish to exclude cases where

  $$A_i = \begin{pmatrix} B_i & D_i \\ & C_i \end{pmatrix} \tag{4.8}$$

  is an upper block-triangular matrix.

- Example.

  $$A_\uparrow = \begin{pmatrix} & 1 \\ & \end{pmatrix}, \quad A_\downarrow = \begin{pmatrix} & \\ 1 & \end{pmatrix}. \tag{4.9}$$

  In this case,

  $$|\psi\rangle = |\uparrow\downarrow\uparrow\downarrow \cdots\rangle + |\downarrow\uparrow\downarrow\uparrow \cdots\rangle , \tag{4.10}$$

  which is a superposition of macroscopically non–translationally invariant states whose translational invariance is restored only after taking a linear combination. As we will see, this can be excluded by imposing constraints on the eigenstructure of the transfer matrix.

We thus wish to exclude states of the above types. For a TI-MPS $\{A_i\}_i$, define the transfer maps

$$T_A(X) := \sum_{i=1}^d A_i X A_i^\dagger, \tag{4.11}$$

$$T_A^*(X) := \sum_{i=1}^d A_i^\dagger X A_i. \tag{4.12}$$

By imposing suitable conditions on these transfer maps, we will obtain the desired class of states.

Proposition 4.3. *The transfer maps T_A, T_A^* are positive linear maps.*

(Proof) Linearity is obvious. For positivity, decompose $X \geq 0$ spectrally as $X = \sum_a \lambda_a P_a$, then

$$T_A(X) = \sum_{i,a} A_i P_a A_i^\dagger, \tag{4.13}$$

and since $A_i P_a A_i^\dagger \geq 0$ (Proposition 3.3), we obtain $T(X) \geq 0$. □

In fact, T_A is known to be completely positive. In what follows, however, we only use the positivity of T.

We now introduce the canonical form of TI-MPS. Let $\{A_i\}_i$ be a TI-MPS. We say that $\{A_i\}_i$ is irreducible (resp. reducible) if T_A is irreducible (resp. reducible).

Proposition 4.4. *If $\{A_i\}_i$ is reducible, then without changing the state $|\psi\rangle$ the TI-MPS can be block-diagonalized as*

$$A_i = \begin{pmatrix} B_i & \\ & C_i \end{pmatrix}. \tag{4.14}$$

(Proof) By reducibility of T_A, there exists a nontrivial orthogonal projection $p \neq 0, 1$ such that $T_A(pMp) \subset pMp$. The restriction of T_A to the subalgebra pMp is also positive, hence there exists $Y \in pMp$ such that $T_A(Y) = \rho_{T_A|_{pMp}} Y$ with $Y \geq 0$, nonzero, and not full rank. With the spectral decomposition

$$Y = \sum_{a=1}^{r} \lambda_a |a\rangle \langle a|, \tag{4.15}$$

we obtain

$$\sum_{a=1}^{r} \sum_{i=1}^{d} A_i |a\rangle \langle a| A_i^\dagger = \sum_{a=1}^{r} \lambda_a |a\rangle \langle a|. \tag{4.16}$$

By general arguments, the vector space spanned by $\{A_i |a\rangle\}_{i,a}$ coincides with that spanned by $\{|a\rangle\}_a$. Let $P_Y = \sum_{a=1}^{r} |a\rangle \langle a|$ be the orthogonal projection onto $\operatorname{im} Y$. Since $A_i |a\rangle \in \operatorname{im} Y$, we have $A_i |a\rangle = P_Y A_i |a\rangle$, hence $A_i P_Y = P_Y A_i P_Y$. Thus, in a basis with $P_Y = \begin{pmatrix} 1_r & \\ & 0 \end{pmatrix}$,

$$A_i = \begin{pmatrix} B_i & D_i \\ 0 & C_i \end{pmatrix}. \tag{4.17}$$

The off-diagonal block does not contribute after taking the trace, so replacing

$$A_i \mapsto \begin{pmatrix} B_i & 0 \\ 0 & C_i \end{pmatrix} \tag{4.18}$$

leaves the state $|\psi\rangle$ unchanged. □

Proposition 4.5. *The similarity transformation (3.97) leaves the state $|\psi\rangle$ unchanged.*

(Proof) It corresponds to the transformation

$$A_i \mapsto \sqrt{c} Q^{-1} A_i Q. \tag{4.19}$$

Here Q cancels out, and $\sqrt{c}$ only contributes an overall constant. □

Combining the above two propositions, we obtain the canonical form of TI-MPS:

Theorem 4.6 ([2]. Canonical form of TI-MPS)**.** *Any TI-MPS $\{A_i\}_i$ admits the canonical form*

$$A_i = \bigoplus_\alpha \rho_\alpha A_i^\alpha, \quad \rho_\alpha > 0, \tag{4.20}$$

where each block $\{A_i^\alpha\}_i$ is irreducible and satisfies the canonical conditions

$$\sum_i A_i^\alpha A_i^{\alpha\dagger} = 1_{D_\alpha}, \tag{4.21}$$

$$\sum_i A_i^{\alpha\dagger} \Lambda^\alpha A_i^\alpha = \Lambda^\alpha, \quad \Lambda^\alpha = \mathrm{diag}(\lambda_1^\alpha, \ldots, \lambda_{D_\alpha}^\alpha), \quad \mathrm{tr}\,[\Lambda^\alpha] = 1. \tag{4.22}$$

4.2 Injectivity and Strong Irreducibility

Irreducibility alone is insufficient to characterize the class of "gapped, nondegenerate ground states." For instance, the undesirable example (4.9) has transfer matrix spectrum $\mathrm{Sp}(T_A) = \{1, -1, 0, 0\}$.

From the behavior of correlation functions, one can impose restrictions on the spectral structure of the transfer matrix T_A. Indeed, it is known that in a gapped ground state, any two-point correlation function decays exponentially [11]. For any pair of operators $\hat{O}_x, \hat{O}'_y$ supported near sites x, y, the correlation function

$$\langle\psi|\hat{O}_x \hat{O}'_y|\psi\rangle \tag{4.23}$$

is governed, in the large-distance limit $|y-x| \to \infty$, by the subleading eigenvalue of $(T_A)^{|y-x|}$. Thus, if there are multiple eigenvalues with $|\lambda| = \rho_{T_A}$, some correlation functions exhibit long-range order. We therefore exclude such cases.

Definition 4.7 (Strong irreducibility [8])**.** A TI-MPS $\{A_i\}_i$ is called strongly irreducible if it is irreducible and

$$|\lambda| = \rho_{T_A} \;\Rightarrow\; \lambda = \rho_{T_A}.$$

If $\{A_i\}_i$ is strongly irreducible, then T_A^N "collapses" for sufficiently large N. With the spectral decomposition of T_A,

$$T_A(Z) = \rho_{T_A} Z, \quad T_A^*(Z^*) = \rho_{T_A} Z^*, \quad \mathrm{tr}\,[Z^* Z] = 1,$$

we can write

$$T_A(\cdot) = \rho_{T_A}\Big[Z\mathrm{tr}\,(Z^* \cdot) + \sum_\lambda \frac{\lambda}{\rho_{T_A}} P_\lambda(\cdot)\Big], \tag{4.24}$$

so that

$$T_A^N \sim (\rho_{T_A})^N \left[Z\mathrm{tr}\,(Z^* \cdot) + \cdots\right]. \tag{4.25}$$

The subleading terms decay exponentially in N. In particular, as $N \to \infty$,

$$T_A^N(X) \sim (\rho_{T_A})^N \mathrm{tr}\,(Z^* X). \tag{4.26}$$

Strong irreducibility can be reformulated as an algebraic property of $\{A_i\}_i$. For $L \in \mathbb{N}$ define the linear map

$$\Gamma_L : \mathrm{Mat}_n(\mathbb{C}) \to (\mathbb{C}^d)^{\otimes L}, \quad X \mapsto \Gamma_L(X) := \sum_{i_1,\dots,i_L} \mathrm{tr}\,[X A_{i_1} \cdots A_{i_L}]\, |i_1 \cdots i_L\rangle\,. \tag{4.27}$$

Proposition 4.8 ([2]).

$$\Gamma_L \textit{ is injective} \quad \Leftrightarrow \quad \mathrm{Span}\,\{A_{i_1} \cdots A_{i_L}\}_{i_1,\dots,i_L} = \mathrm{Mat}_n(\mathbb{C}).$$

(Proof) ($\Leftarrow$) If $\Gamma_L(X) = 0$, then $\mathrm{tr}\,[X A_{i_1} \cdots A_{i_L}] = 0$ for all $i_1, \dots, i_L$. By assumption, $\{A_{i_1} \cdots A_{i_L}\}$ spans $\mathrm{Mat}_n(\mathbb{C})$, so there exist coefficients $\{c_{i_1 \dots i_L}\}$ with $\sum c_{i_1 \dots i_L} A_{i_1} \cdots A_{i_L} = X^\dagger$. Then $\mathrm{tr}\,[XX^\dagger] = 0$, i.e. $X = 0$.

($\Rightarrow$) Represent Γ_L as

$$[\Gamma_L]_{i_1 \dots i_L, ab} = \langle i_1 \cdots i_L | \Gamma_L(|a\rangle\langle b|)\rangle = \langle b| A_{i_1} \cdots A_{i_L} |a\rangle = [A_{i_1} \cdots A_{i_L}]_{ba}. \tag{4.28}$$

If Γ_L is injective, the matrix Γ_L has full rank n^2, hence the row vectors $[A_{i_1} \cdots A_{i_L}]$ span $\mathrm{Mat}_n(\mathbb{C})$. □

Proposition 4.9 ([2]).

$$\Gamma_L \textit{ injective} \quad \Rightarrow \quad \Gamma_{L'} \textit{ injective for all } L' > L.$$

(Proof) For $L' = L + 1$,

$$\Gamma_{L+1}(X) = \sum_{i_{L+1}} \Gamma_L(A_{i_{L+1}} X)\, |i_{L+1}\rangle\,.$$

If $\Gamma_{L+1}(X) = 0$, then $\Gamma_L(A_i X) = 0$ for all i. By injectivity of Γ_L, this implies $A_i X = 0$ for all i. Since $\{A_{i_1} \cdots A_{i_L}\}$ spans $\mathrm{Mat}_n(\mathbb{C})$, there exist $\{c_{i_1 \dots i_L}\}$ with $\sum c_{i_1 \dots i_L} A_{i_1} \cdots A_{i_L} = 1_n$. Thus $X = 0$. □

Definition 4.10 (Injectivity)**.** A TI-MPS $\{A_i\}_i$ is injective if there exists $L_0 \in \mathbb{N}$ such that Γ_{L_0} is injective. Equivalently, for some L_0, the set $\{A_{i_1} \cdots A_{i_{L_0}}\}_{i_1,\dots,i_{L_0}}$ spans $\mathrm{Mat}_n(\mathbb{C})$.

Note that injectivity of $\{A_i\}_i$ is different from requiring that the set of all products of arbitrary length spans $\mathrm{Mat}_n(\mathbb{C})$, i.e.

$$\mathrm{Span}\left(\bigcup_L \{A_{i_1} \cdots A_{i_L}\}_{i_1,\dots,i_L}\right) = \mathrm{Mat}_n(\mathbb{C}).$$

For example, in (4.9), although $\mathrm{Span}(A_\uparrow, A_\downarrow, A_\uparrow A_\downarrow, A_\downarrow A_\uparrow) = \mathrm{Mat}_2(\mathbb{C})$, for no fixed L do the length-L products span $\mathrm{Mat}_2(\mathbb{C})$.

Proposition 4.11. *If $\{A_i\}_i$ is injective, then it is irreducible.*

(Proof) By contraposition. If $\{A_i\}_i$ is reducible, then as in Proposition 4.4,

$$A_i = \begin{pmatrix} B_i & D_i \\ 0 & C_i \end{pmatrix},$$

so for any L, the products $\{A_{i_1} \cdots A_{i_L}\}$ cannot span $\mathrm{Mat}_n(\mathbb{C})$. □

Proposition 4.12 ([1, 8]). *If $\{A_i\}_i$ is strongly irreducible, then it is injective.*

(Proof) Suppose not: for every L, the products $\{A_{i_1} \cdots A_{i_L}\}$ do not span $\mathrm{Mat}_n(\mathbb{C})$. Then for each L there exists $B_L \neq 0$ orthogonal to this span, i.e. $\mathrm{tr}\,[A_{i_1} \cdots A_{i_L} B_L] = 0$ for all $i_1, \ldots, i_L$. Thus

$$0 = (T_A)^L (B_L \otimes B_L^\dagger).$$

But by strong irreducibility,

$$(T_A)^L (B_L \otimes B_L^\dagger) \xrightarrow{L \to \infty} (\rho_{T_A})^L \mathrm{tr}\,(Z^* B_L Z B_L^\dagger),$$

which is nonzero since $Z, Z^* > 0$. Contradiction. □

In fact, the converse also holds: injectivity implies strong irreducibility [8]. The proof relies on spectral properties of T_A following from Schwarz-positivity [7], which we do not reproduce here.

4.3 Fundamental theorem of TI-MPS

For a finite system, the quantum state $|\psi\rangle$ is defined only up to a global $U(1)$ phase, i.e. $|\psi\rangle \sim e^{i\alpha} |\psi\rangle$. Similarly, a TI-MPS $\{A_i\}_i$ leaves the state $|\psi\rangle$ invariant (up to an overall factor) under the similarity transformation (4.19); that is, the "physical state" is unchanged. While a general TI-MPS admits more ambiguities than (4.19), in the strongly irreducible case it can be shown that the only freedom is precisely the similarity transformation (4.19).

We first relate the bond dimension n of a TI-MPS to the structure of the state $|\psi\rangle$ it represents.

Proposition 4.13 ([2]). *Let $\{A_i\}_i$ be an injective TI-MPS with bond dimension n, and suppose Γ_{L_0} is injective for some $L_0 \in \mathbb{N}$. For a system of N sites and any R with $L_0 \leq R \leq N - R$, one has*

$$\mathrm{rank}\, \rho_{[1,R]} = n^2.$$

(Proof) We can write

$$|\psi\rangle = \sum_{i_1 \cdots i_N} \mathrm{tr}\,[A_{i_1} \cdots A_{i_N}] \, |i_1 \cdots i_N\rangle \tag{4.29}$$

$$= \sum_{a,b=1}^{n} \Gamma_R(|b\rangle \langle a|) \, \Gamma_{N-R}(|a\rangle \langle b|). \tag{4.30}$$

By injectivity of Γ_R and Γ_{N-R}, the n^2 vectors $\{\Gamma_R(|b\rangle \langle a|)\}_{a,b}$ and $\{\Gamma_{N-R}(|a\rangle \langle b|)\}_{a,b}$ are linearly independent. Thus

$$|\psi\rangle = \sum_{k=1}^{n^2} |v_k\rangle \, |w_k\rangle \,, \quad \{|v_k\rangle\}_k, \{|w_k\rangle\}_k \text{ linearly independent.}$$

Hence

$$\rho_{[1,R]} = \sum_{k,l=1}^{n^2} |v_k\rangle \langle w_l | w_k \rangle \langle v_l| \,.$$

Since the Gram matrix $\langle w_l | w_k \rangle$ is full rank, $\mathrm{rank}(\rho_{[1,R]}) = n^2$. □

We now introduce two technical lemmas.

Lemma 4.14 ([2]). *Let $T, S : \mathbb{C}^\ell \to \mathbb{C}^m$ be linear maps and let $Y_1, \dots, Y_\ell \in \mathbb{C}^\ell$ satisfy:*

- $T(Y_k) = S(Y_{k+1})$ *for* $k = 1, \dots, \ell - 1$.
- $Y_1, \dots, Y_{\ell-1}$ *are linearly independent.*
- $Y_1, \dots, Y_\ell$ *are linearly dependent, with* $Y_\ell = \sum_{k=1}^{\ell-1} \lambda_k Y_k$.

Then for any solution x of the polynomial equation

$$\lambda_1 x^{\ell-1} + \lambda_2 x^{\ell-2} + \cdots + \lambda_{\ell-1} x = 1,$$

define

$$\mu_1 = \lambda_1 x, \tag{4.31}$$
$$\mu_2 = \lambda_1 x^2 + \lambda_2 x, \tag{4.32}$$
$$\vdots \tag{4.33}$$
$$\mu_{\ell-1} = \lambda_1 x^{\ell-1} + \cdots + \lambda_{\ell-1} x = 1. \tag{4.34}$$

Then the vector

$$Y = \sum_{k=1}^{\ell-1} \mu_k Y_k$$

satisfies

$$Y \neq 0, \quad T(Y) = \frac{1}{x} S(Y).$$

(Proof omitted; direct calculation.)

Lemma 4.15 ([2]). *Let $B, C \in \mathrm{Mat}_n(\mathbb{C})$. The solution space of*

$$W(C \otimes 1_n) = (B \otimes 1_n)W$$

is

$$S \otimes \mathrm{Mat}_n(\mathbb{C}), \quad S = \{X \in \mathrm{Mat}_n(\mathbb{C}) \mid XC = BX\}.$$

(Proof omitted.)

Theorem 4.16 ([2, 12]. Fundamental theorem of TI-MPS). *Let $\{B_i\}_i$ be a TI-MPS of bond dimension n, which is injective and in the canonical form, that is, there exists $L_0 \in \mathbb{N}$ such that Γ_{L_0} is injective and the canonical conditions*

$$\sum_i B_i B_i^\dagger = 1_n, \tag{4.35}$$

$$\sum_i B_i^\dagger \Lambda_B B_i = \Lambda_B, \quad \Lambda_B = \mathrm{diag}(\lambda_{B,1}, \ldots, \lambda_{B,n}), \quad \lambda_{B,1} \geq \cdots \geq \lambda_{B,n} > 0, \quad \mathrm{tr}\,[\Lambda_B] = 1, \tag{4.36}$$

are satisfied. Furthermore, let $\{C_i\}_i$ be another TI-MPS of the same bond dimension n, which is irreducible and in the canonical form, that is, satisfying

$$\sum_i C_i C_i^\dagger = 1_n, \tag{4.37}$$

$$\sum_i C_i^\dagger \Lambda_C C_i = \Lambda_C, \quad \Lambda_C = \mathrm{diag}(\lambda_{C,1}, \ldots, \lambda_{C,n}), \quad \lambda_{C,1} \geq \cdots \geq \lambda_{C,n} > 0, \quad \mathrm{tr}\,[\Lambda_C] = 1. \tag{4.38}$$

Suppose that for some system size $N > 2L_0 + n^4$, the states represented by the TI-MPS $\{B_i\}_i$ and $\{C_i\}_i$ are physically equivalent, that is, there exists a $U(1)$ phase $e^{i\alpha}$ such that

$$\psi_{i_1 \cdots i_N} = \mathrm{tr}\,[B_{i_1} \cdots B_{i_N}] = e^{i\alpha} \mathrm{tr}\,[C_{i_1} \cdots C_{i_N}] \quad \textit{for all } i_1, \ldots, i_N. \tag{4.39}$$

Then, there exist a unitary matrix $U \in U(n)$ and a $U(1)$ phase $e^{i\theta} \in U(1)$ such that

$$B_i = e^{i\theta} U C_i U^\dagger, \quad i = 1, \ldots, d. \tag{4.40}$$

Moreover, if the TI-MPS $\{B_i\}_i, \{C_i\}_i$ are strongly irreducible, then $e^{i\theta}$ is unique and U is unique up to a $U(1)$ phase.

The strategy of the proof is to construct OBC-MPS from TI-MPS $\{B_i\}, \{C_i\}_i$, and then use the uniqueness of OBC-MPS (Proposition 2.2) to relate them.

(Proof) By redefining $C_i \mapsto e^{-i\alpha/N} C_i$, we may assume $e^{i\alpha} = 1$. Construct OBC-MPS as follows:

$$\mathrm{tr}\,[B_{i_1} \cdots B_{i_N}] = \sum_{a,b,c} [B_{i_1}]_{ab} [B_{i_2} \cdots B_{i_{N-1}}]_{bc} [B_{i_N}]_{ca} = \sum_{a,b,c} [B_{i_1}]_{ab} [\tilde{B}]_{bc} [B_{i_N}]_{ca} \tag{4.41}$$

where, regarding the index a as contracted by an inner product, we arrange $\tilde{B}$ diagonally n times as

$$= \sum_a u_a \tilde{B} v_a \tag{4.42}$$

$$= (u_1, \ldots, u_n) \begin{pmatrix} \tilde{B} & & \\ & \ddots & \\ & & \tilde{B} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. \tag{4.43}$$

Define

$$b_i^{[1]} = (u_1, \ldots, u_n) = ([B_i]_{1\cdot}, \ldots, [B_i]_{n\cdot}), \tag{4.44}$$

$$b_i^{[N]} = (v_1, \ldots, v_n)^\top = ([B_i]_{\cdot 1}, \ldots, [B]_{\cdot n})^\top, \tag{4.45}$$

then we obtain the OBC-MPS representation

$$\psi_{i_1 \cdots i_N} = b_{i_1}^{[1]} (B_{i_2} \otimes 1_n) \cdots (B_{i_{N-1}} \otimes 1_n) b_{i_N}^{[N]}. \tag{4.46}$$

Similarly for $\{C_i\}_i$,

$$\psi_{i_1 \cdots i_N} = c_{i_1}^{[1]} (C_{i_2} \otimes 1_n) \cdots (C_{i_{N-1}} \otimes 1_n) c_{i_N}^{[N]}. \tag{4.47}$$

By the canonical form of OBC-MPS (Proposition 2.2), there exists a **common** standard OBC-MPS $\{\{A^{[m]}_{i_m}\}_{i_m}\}_m$ such that

$$A^{[m]}_i = Y_{m-1}(B_i \otimes 1_n)Z_m = Y'_{m-1}(C_i \otimes 1_n)Z'_m, \tag{4.48}$$
$$Y_m Z_m = 1, \quad Y'_m Z'_m = 1, \quad 1 < m < N, \tag{4.49}$$

holds. $A^{[m]}_{i_m}$ is an $r_{m-1} \times r_m$ matrix. By Proposition 4.13, for $L_0 \leq m \leq N - L_0$, the reduced density matrix $\rho_{[1,m]}$ has rank n^2. Thus, for $L_0 \leq m \leq N - L_0$, $r_m = n^2$. The matrices Y_m, Y'_m are $r_m \times n^2$, and Z_m, Z'_m are $n^2 \times r_m$, so (4.49) implies that Y_m, Z_m, Y'_m, Z'_m are invertible for $L_0 \leq m \leq N - L_0$. Therefore, from (4.48), defining $W_m = Z_m Y'_m$, we have

$$W_{m-1}(C_i \otimes 1_n) = (B_i \otimes 1_n)W_m, \quad i = 1, \dots, d, \quad L_0 + 1 \leq m \leq N - L_0, \tag{4.50}$$

which holds independently of i. The set $\{W_m\}_{m=L_0+1}^{N-L_0}$ is a solution to the linear equations for the $n^2 \times n^2$ matrices, and since $N - 2L_0 > n^4$, they are linearly dependent.

Thus, we have obtained the following: define $T_i(W) = W(B_i \otimes 1_n)$, $S_i(W) = (C_i \otimes 1_n)W$. There exists $\ell \in \mathbb{N}, 1 < \ell < n^4 + 1$ such that $W_1, \dots, W_{\ell-1}$ are linearly independent and

$$W = W_\ell = \sum_{k=1}^{\ell-1} \lambda_k W_k, \tag{4.51}$$
$$T_i(W_k) = S_i(W_{k+1}), \quad k = 1, \dots, \ell - 1, \quad i = 1, \dots, d. \tag{4.52}$$

Therefore, by Lemma 4.14, there exist $W \neq 0$ and $x \neq 0$ such that

$$W(B_i \otimes 1_n) = \frac{1}{x}(C_i \otimes 1_n)W, \quad i = 1, \dots, n. \tag{4.53}$$

By Lemma 4.15, W belongs to the following intersection:

$$W \in \bigcap_{i=1}^{d} S_i \otimes \mathrm{Mat}_n(\mathbb{C}) = (\bigcap_{i=1}^{d} S_i) \otimes \mathrm{Mat}_n(\mathbb{C}), \quad S_i = \{X \in \mathrm{Mat}_n(\mathbb{C}) \mid XB_i = \frac{1}{x}C_i X\}. \tag{4.54}$$

In particular, $\bigcap_{i=1}^{d} S_i \neq \emptyset$. Choose $U \in \bigcap_{i=1}^{d} S_i$, then

$$UB_i = \frac{1}{x}C_i U, \quad i = 1, \dots, d. \tag{4.55}$$

We show that $|x| = 1$. Since $\{B_i\}_i$ is in canonical form,

$$U^\dagger \Lambda_B U = U\Big(\sum_i B_i^\dagger \Lambda_B B_i\Big)U^\dagger = \frac{1}{|x|^2}\sum_i C_i^\dagger U^\dagger \Lambda_B U C_i. \tag{4.56}$$

Taking the trace, we obtain

$$0 < \mathrm{tr}\,[U^\dagger \Lambda_B U] = \frac{1}{|x|^2}\sum_i \mathrm{tr}\,[U^\dagger \Lambda_B U C_i C_i^\dagger] = \frac{1}{|x|^2}\mathrm{tr}\,[U^\dagger \Lambda_B U], \tag{4.57}$$

hence $|x| = 1$.

We show that $U \in U(n)$. From $UB_i = e^{i\theta}C_i U$ we obtain

$$T_C(UU^\dagger) = \sum_i C_i U U^\dagger C_i^\dagger = \sum_i U B_i B_i^\dagger U^\dagger = UU^\dagger. \tag{4.58}$$

Since T_C is irreducible, the eigenvector corresponding to $\lambda = 1$ is unique. Therefore, $UU^\dagger \propto 1_n$. As $UU^\dagger > 0$, we may set $UU^\dagger = 1_n$.

We now prove the uniqueness of $e^{i\theta}, U$ under the assumption of strong irreducibility. Suppose $U'B_i = e^{i\theta'}C_iU,\ i = 1, \ldots, d$.

$$\sum_i C_i U' U^\dagger C_i^\dagger = \sum_i e^{-i\theta'} U' B_i e^{i\theta} B_i^\dagger U^\dagger = e^{-i\theta'} e^{i\theta} U' U^\dagger. \tag{4.59}$$

By strong irreducibility, the only eigenvalue λ with $|\lambda| = 1$ is $\lambda = 1$, and the corresponding eigenvector is unique. Therefore, $e^{i\theta'} = e^{i\theta}$ and $U' \sim U$. □

As already noted, injectivity and strong irreducibility are equivalent [8], so it is not necessary to impose strong irreducibility additionally to prove the uniqueness of $e^{i\theta}, U$.

4.4 Application: one-dimensional SPT phases

As an application, we show that when a translationally invariant and "gapped non-degenerate" state $|\psi\rangle$ has a symmetry group G, an invariant taking values in the group cohomology $H^1(G, U(1)_\phi) \times H^2(G, U(1)_\phi)$ can be defined.

First, let us define the symmetry. Let G be a group, and let $\phi : G \to \mathbb{Z}/2 = \{\pm 1\}$ be a homomorphism such that $\phi_g = 1$ indicates that g is a unitary symmetry, while $\phi_g = -1$ indicates that g is an antiunitary symmetry. For a group element $g \in G$, its action on the Hilbert space $\mathcal{H}$ is defined as the tensor product of the group action on the local Hilbert spaces:

$$\hat{g} = \begin{cases} \bigotimes_x \hat{g}_x & (\phi_g = 1), \\ (\bigotimes_x \hat{g}_x) K & (\phi_g = -1), \end{cases} \quad \hat{g}_x |j\rangle = \sum_i |i\rangle [u_g]_{ij}. \tag{4.60}$$

Here K denotes complex conjugation. The unitary matrices $\{u_g\}_g$ are assumed to form a linear representation, that is,

$$u_g u_h^{\phi_g} = u_{gh}, \quad g, h \in G, \tag{4.61}$$

where for a matrix X we introduce the notation

$$X^{\phi_g} = \begin{cases} X & (\phi_g = 1), \\ X^* & (\phi_g = -1). \end{cases} \tag{4.62}$$

Writing the state described by the TI-MPS $\{A_i\}_i$ as

$$|\{A_i\}_i\rangle = \sum_{i_1, \ldots, i_N} \mathrm{tr}\left[A_{i_1} \cdots A_{i_N}\right] |i_1 \cdots i_N\rangle, \tag{4.63}$$

the action of G on the TI-MPS is given by

$$\hat{g} |\{A_i\}_i\rangle = \sum_{i_1, \ldots, i_N, j_1, \ldots, j_N} \mathrm{tr}\left[A_{i_1} \cdots A_{i_N}\right]^{\phi_g} |j_1 \cdots j_N\rangle [u_g]_{j_1 i_1} \cdots [u_g]_{j_N i_N} = \left| \left\{ \sum_j [u_g]_{ij} A_j^{\phi_g} \right\} \right\rangle, \tag{4.64}$$

that is,

$$g : A_i \mapsto {}^g A_i := \sum_j [u_g]_{ij} A_j^{\phi_g}, \quad i = 1, \ldots, d. \tag{4.65}$$

By the linearity of u_g, note that

$${}^{gh} A_i = {}^{gh} A_i, \quad i = 1, \ldots, d. \tag{4.66}$$

The state $|\{A_i\}_i\rangle$ is G-invariant, i.e., there exists $e^{i\alpha_g}$ such that

$$\hat{g} |\{A_i\}_i\rangle = |\{{}^g A_i\}_i\rangle = e^{i\alpha_g} |\{A_i\}_i\rangle, \quad g \in G, \tag{4.67}$$

where the system size N is fixed.

It is expected that any translationally invariant "gapped non-degenerate state" can be approximated by a strongly irreducible TI-MPS. Let $\{A_i\}_i$ be a strongly irreducible TI-MPS in canonical form. Then, by Theorem 4.16, the G symmetry (4.67) implies, for each $g \in G$,

$$ {}^{g}A_i = e^{i\theta_g} V_g^{\dagger} A_i V_g, \quad i = 1, \dots, d, \tag{4.68} $$

where $e^{i\theta_g}$ is unique and the unitary matrix V_g is unique up to a $U(1)$ phase. From (4.66) we have

$$ \begin{aligned} {}^{gh}A_i &= {}^{g}(e^{i\theta_h} V_h^{\dagger} A_i V_h) = e^{i\phi_g\theta_h} (V_h^{\phi_g})^{\dagger} ({}^{g}A_i) V_h^{\phi_g} \\ &= e^{i\phi_g\theta_h} (V_h^{\phi_g})^{\dagger} e^{i\theta_g} V_g^{\dagger} A_i V_g V_h^{\phi_g} = e^{i\theta_g} e^{i\phi_g\theta_h} (V_g V_h^{\phi_g})^{\dagger} A_i (V_g V_h^{\phi_g}), \quad i = 1, \dots, d. \end{aligned} \tag{4.69} $$

On the other hand,

$$ {}^{gh}A_i = e^{i\theta_{gh}} V_{gh}^{\dagger} A_i V_{gh}. \tag{4.70} $$

By uniqueness, $\{e^{i\theta_g}\}_g$ forms a one-dimensional representation of G, that is, $\{e^{i\theta_g}\}_g \in \mathrm{Hom}(G, U(1)_\phi) \cong H^1(G, U(1)_\phi)$:

$$ e^{i\theta_g} e^{i\phi_g\theta_h} = e^{i\theta_{gh}}, \quad g, h \in G. \tag{4.71} $$

On the other hand, since V_g is unique up to a $U(1)$ phase, there exists $z_{g,h} \in U(1)$ such that

$$ V_g V_h^{\phi_g} = z_{g,h} V_{gh}, \quad g, h \in G. \tag{4.72} $$

The $U(1)$ phase $z_{g,h}$ is not arbitrary, and from $(V_g V_h^{\phi_g}) V_k^{\phi_{gh}} = V_g (V_h V_k^{\phi_h})^{\phi_g}$ we obtain

$$ z_{h,k}^{\phi_g} z_{gh,k}^{-1} z_{g,hk} z_{g,h}^{-1} = 1, \quad g, h, k \in G. \tag{4.73} $$

That is, $z_{g,h}$ satisfies the 2-cocycle condition. In other words, $z \in Z^2(G, U(1)_\phi)$ [11]. A change of the $U(1)$ phase of V_g, $V_g \mapsto V_g \eta_g$, induces the transformation

$$ z_{g,h} \mapsto z_{g,h} \eta_h^{\phi_g} \eta_{gh}^{-1} \eta_g, \quad g, h \in G, \tag{4.74} $$

where $\{\eta_g\}_g$ is an element of the group of 2-coboundaries $B^2(G, U(1)_\phi)$. Therefore, a strongly irreducible, canonical, and G-invariant TI-MPS $\{A_i\}_i$ determines an element of the second cohomology $H^2(G, U(1)_\phi) = Z^2(G, U(1)_\phi)/B^2(G, U(1)_\phi)$.

Thus, we obtain the following:

Theorem 4.17 ([13, 14], one-dimensional SPT phases). *Let $\{A_i\}_i$ be a strongly irreducible, canonical, and G-invariant TI-MPS. Then $\{A_i\}_i$ defines quantities taking values in*

$$ H^1(G, U(1)_\phi) \times H^2(G, U(1)_\phi). \tag{4.75} $$

As a matter of fact, constructions of TI-MPS realizing all possible combinations of invariants in $H^1(G, U(1)_\phi) \times H^2(G, U(1)_\phi)$ are known [15].

References

[1] Mark Fannes, Bruno Nachtergaele, and Reinhard F. Werner, *Finitely correlated states on quantum spin chains*, Communications in mathematical physics **144** (1992): 443-490.

[11] $U(1)_\phi$ means that the G action is defined as $g.z = z^{\phi_g}$, $z \in U(1)$.

[2] D. Perez-Garcia, F. Verstraete, M.M. Wolf, J.I. Cirac, *Matrix Product State Representations*, arXiv:quant-ph/0608197.

[3] Ignacio Cirac, David Perez-Garcia, Norbert Schuch, Frank Verstraete *Matrix Product States and Projected Entangled Pair States: Concepts, Symmetries, and Theorems*, arXiv:2011.12127.

[4] Ulrich Schollwoeck, *The density-matrix renormalization group in the age of matrix product states*, arXiv:1008.3477.

[5] Michael M. Wolf, *Quantum Channels & Operations, Guided Tour*, lecture note, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=afa58291b0b8bd47504acb1ab8f553f0b37685cf

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

[7] Evans, David E. and Raphael Hoegh-Krohn, *Spectral properties of positive maps on C*-algebras.* Preprint series: Pure mathematics http://urn. nb. no/URN: NBN: no-8076 (1977).

[8] Sanz, M., Perez-Garcia, D., Wolf, M. M., Cirac, J. I. (2010). *A quantum version of Wielandt's inequality*, IEEE Transactions on Information Theory, 56(9), 4668-4673.

[9] Bhatia, R. (2009). *Positive definite matrices*, Princeton university press.

[10] H. H. Schaefer, M. P. Wolff, Topological Vector Spaces, Springer-Verlag New York, Inc. 1999.

[11] Matthew B. Hastings, Tohru Koma, *Spectral Gap and Exponential Decay of Correlations*, arXiv:math-ph/0507008.

[12] D. Perez-Garcia, M.M. Wolf, M. Sanz, F. Verstraete, J.I. Cirac *String order and symmetries in quantum spin lattices*, arXiv:0802.0447.

[13] Frank Pollmann, Erez Berg, Ari M. Turner, Masaki Oshikawa *Entanglement spectrum of a topological phase in one dimension*, arXiv:0910.1811

[14] Xie Chen, Zheng-Cheng Gu, Xiao-Gang Wen, *Classification of Gapped Symmetric Phases in 1D Spin Systems*, arXiv:1008.3745.

[15] Xie Chen, Zheng-Cheng Gu, Zheng-Xin Liu, Xiao-Gang Wen, *Symmetry protected topological orders and the group cohomology of their symmetry group*, arXiv:1106.4772.