Thermal State Preperation with Petz Recovery

Wenhao He

April 11, 2025

1 INTRODUCTION

Preparing quantum thermal states has broad application. In this note, we introduce a thermalization strategy that leverages the Petz recovery map. Central to our approach is the capability of the Petz recovery map to form a composite map admitting a predetermined fixed point. This property allows us to construct a projector targeting this fixed point by repeatedly applying the composite map. This methodology presents a novel approach to prepare thermal or ground quantum states, showcasing the practical utility of the Petz recovery map in quantum computing applications.

1.1 Related literature

Chen et. al designed a Lidbladian with the target as its steady state to do the thermalization [CKG23][CKBG23]. And Ding et. al generalized this framework[DLL24]. Their gate complexity is only logarithmically on the precision. Bakshi et. al pointed out that high-temperature Gibbs state can be efficiently prepared [BLMT24]. Zhang et. al proposed dissipative algorithms [MZAW24] [ZBC23]. Bergamaschi et. al proved quantum advantage of Gibbs sampling [BCL24].

Possible applications on material science [AAB+24].

2 BACKGROUND

2.1 Problem settings

The basic problem is to construct a Gibbs state for a given Hamiltonian. We formulate it in Problem 1. In many applications we require $\hat{H}$ to be lattice model and assume it to be k-local, but in this note(at current stage) we neglect such assumptions and consider general $\hat{H}$ for simplicity. In this note it is very important to distinguish between quantum states and quantum operations, so we add hat for the latter and no hat for the former. We use mathcal letters to represent channels or superoperators.

Problem 1 (Thermal state preparation)**.** *Given a Hamiltonian $\hat{H}$ and inverse of temperature β, output the quantum Gibbs state*

$$\sigma_H = \frac{e^{-\beta \hat{H}}}{\operatorname{Tr} e^{-\beta \hat{H}}} \tag{1}$$

2.2 Petz recovery

Petz recovery is a quantum analog of Bayes rule. For any quantum channel $\mathcal{E} : H_A \longrightarrow H_B$, Petz recovey provides an approximate 'inverse' map, which reads

$$\mathcal{R}_\sigma^{\mathcal{E}}(\cdot) = \sum_i \sigma^{\frac{1}{2}} \hat{E}_i^\dagger \mathcal{E}(\sigma)^{-\frac{1}{2}} (\cdot) \mathcal{E}(\sigma)^{-\frac{1}{2}} \hat{E}_i \sigma^{\frac{1}{2}} \tag{2}$$

where σ is a predetermined quantum state, $^{-\frac{1}{2}}$ is pseudo inverse, and $\{\hat{E}_i\}$ is Kraus representation for channel $\mathcal{E}$. We can implement this channel using [GLM+22].

The predetermined quantum state σ is fixed point of the composition map of $\mathcal{E}$ followed by Petz:

$$\mathcal{R}_\sigma^{\mathcal{E}} \circ \mathcal{E}(\sigma) = \sigma. \tag{3}$$

We can verify this relation directly. Here we note that this composition map satisfies the Kubo – Martin – Schwinger (KMS) detailed balance condition, so we can also prove the fixed point property utilizing this KMS condition.

2.3 QEC matrix based calculations

Recently [ZHLJ24] introduces a framework to analyze transpose recovery map(which is a special case of Petz recovery). In the context of quantum error correction, they propose to expand operators on $E_i|\mu\rangle$ basis, where E_is are Kraus operator for noisy channel, and $|\mu\rangle$ is codeword basis. By exploiting this basis, they reduce some important formula to functions barely dependent of QEC matrix. Similar tricks can be used in this note.

3 METHODS&RESULTS

3.1 High level strategy

We plan to break Problem 1 into three parts: sub-Problem 1, sub-Problem 2 and sub-Problem 3. For sub-Problem 1 imaginary time evolution, we provide two strategies: the Linear Combination of Unitaries (LCU) based method (wenhao: possible related literature:[ALL23]) and the Hubbard–Stratonovich transformation (H-S transformation) based method[CS16]. For sub-Problem 2 Petz recovery construction, we also provide two strategies: Quantum Singular Value Transformation (QSVT) [GLM+22] based method and Quantum Polar Decomposition (QPD) based method [QR22]. Then for the last sub-Problem 3 Converging to steady state we analyze the mixing time and discuss how to choose tunable noise channels $\mathcal{E}$. Our roadmap is summerized in Figure 1.

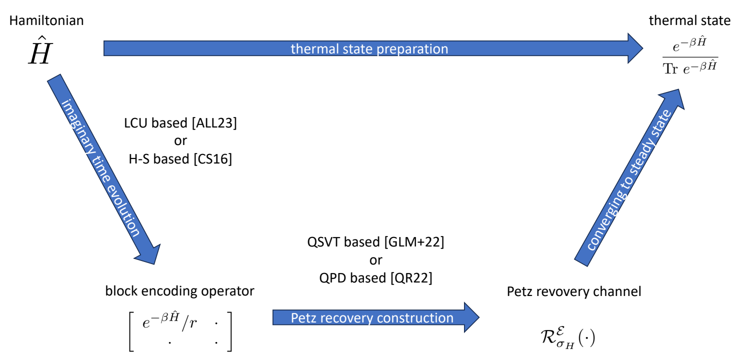


Figure 1: A visualization of our high level strategy.

sub-Problem 1 (Imaginary time evolution). *Given access to a Hamiltonian operator $\hat{H}$ and temperature inverse β, implement block encoding operation of imaginary time evolution*

$$\hat{V}_{\sigma_H} = \begin{bmatrix} e^{-\beta\hat{H}}/r & \cdot \\ \cdot & \cdot \end{bmatrix}. \tag{4}$$

sub-Problem 2 (Petz recovery construction). *Given a block encoding imaginary time evolution operator $\hat{V}_{\sigma_H} = \begin{bmatrix} e^{-\beta\hat{H}}/r & \cdot \\ \cdot & \cdot \end{bmatrix}$ and access to a tunable noise channel $\mathcal{E}$, implement the Petz recovery quantum channel*

$$\mathcal{R}^{\mathcal{E}}_{\sigma_H}(\cdot) = \sum_i \sigma_H^{\frac{1}{2}} \hat{E}_i^\dagger \mathcal{E}(\sigma_H)^{-\frac{1}{2}} (\cdot) \mathcal{E}(\sigma_H)^{-\frac{1}{2}} \hat{E}_i \sigma_H^{\frac{1}{2}}, \tag{5}$$

where $\{\hat{E}_i\}$ is the Kraus representation of noise channel $\mathcal{E}$, and $\sigma_H = e^{-\beta\hat{H}} / \mathrm{Tr}\, e^{-\beta\hat{H}}$ is the thermal state.

sub-Problem 3 (Converging to steady state)**.** *Given access to a noise channel $\mathcal{E}$ and its corresponding Petz recovery channel $\mathcal{R}^{\mathcal{E}}_{\sigma_H}$, prepare the thermal state $\sigma_H = e^{-\beta\hat{H}} / \operatorname{Tr} e^{-\beta\hat{H}}$.*

3.2 Imaginary time evolution with LCU

The motivation of designing an imaginary time evolution operation is the observation that Petz recovery expression Eq. (2) includes a term of σ, which is proportional to $e^{-\beta\hat{H}}$ in our case. This implies that the construction of Petz recovery includes implementing the non-unitary operator $e^{-\beta\hat{H}}$ in some way. Here we emphasize that our task in this subsection is to construct a block encoding operation of $e^{-\beta\hat{H}}$, instead of obtaining a copy of the state σ_H.

Although imaginary time evolution operation is hard to implement in general, we can decompose it into real time evolution, i.e. Hamiltonian simulation, which is well studied and easier to implement. One way to do such decomposition is doing LCU with coefficients obtained by 'Fourier transformation':

$$e^{-\beta\hat{H}} = \int_{-\infty}^{\infty} \frac{1}{\pi(1+k^2)} e^{-ik\beta\hat{H}} dk. \tag{6}$$

With identity Eq. (6) generally true, we want to truncate and discretize it into a finite sum to be compatible with LCU framework:

$$e^{-\beta\hat{H}} = \sum_{l=-\lfloor\frac{k_{max}}{\Delta k}\rfloor}^{\lfloor\frac{k_{max}}{\Delta k}\rfloor} \frac{\Delta k}{\pi(1+(l\Delta k)^2)} e^{-il\Delta k\beta H} + \epsilon_{trucate} + \epsilon_{discrete} \tag{7}$$

$$\tag{8}$$

where the truncation error and discretization error $\epsilon_{trucate}$ and $\epsilon_{discrete}$ evaluate as following:

$$\epsilon_{truncate} = O(\int_{k_{max}}^{\infty} \frac{dk}{1+k^2}) = O(\frac{1}{k_{max}}) \tag{9}$$

$$\epsilon_{discretize} = O(k_{max}\Delta k^2 \max_{|l|\leq\lfloor\frac{k_{max}}{\Delta k}\rfloor} \frac{\partial^2}{\partial(l\Delta k)^2}(\frac{e^{-il\Delta k\beta H}}{1+(l\Delta k)^2})) = O(k_{max}\Delta k^2\beta^2||H||^2), \tag{10}$$

where we can choose

$$k_{max} = O(\frac{1}{\epsilon}) \tag{11}$$

$$\Delta k = (\frac{\epsilon}{\beta||H||}) \tag{12}$$

to keep the total error within $O(\epsilon)$ scale.

There are two sources of gate complexity, one is implementing LCU with finite sum Eq. (7):

$$\sum_{l=-\lfloor\frac{k_{max}}{\Delta k}\rfloor}^{\lfloor\frac{k_{max}}{\Delta k}\rfloor} |\frac{\Delta k}{\pi(1+(l\Delta k)^2)}| = O(1), \tag{13}$$

while the other is the total time of doing Hamiltonian simulation:

$$\tau = \sum_{l=-\lfloor\frac{k_{max}}{\Delta k}\rfloor}^{\lfloor\frac{k_{max}}{\Delta k}\rfloor} l\Delta k\beta = O(\frac{k_{max}^2}{\Delta k}\beta). \tag{14}$$

Substituting parameters in Eqs. (11) (12) and utilizing Hamiltonian simulation [BCS+20], we obtain the overall complexity of implementing imaginary time evolution:

$$N_{\sigma_H} = \tilde{O}(\frac{\beta^2||H||^2}{\epsilon^3}) \tag{15}$$

3.3 Imaginary time evolution with H-S transformation

From above we could observe that complexity of implementing LCU largely depends on the coefficient. Polynomial dependence on $1/k$ of coefficient in Eq. (6) leads to the polynomial dependence on $1/\epsilon$ of gate complexity. We naturally want to find another way to do LCU with more concentrated coefficients. Fortunately, the H-S transformation provides a Gaussian concentration:

$$e^{-\beta\hat{H}} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dy e^{-\frac{y^2}{2}} e^{-iy\sqrt{2\beta\hat{H}}}. \tag{16}$$

To avoid complexity from implementing $\sqrt{\hat{H}}$, we can construct a effective square root Hamiltonian $\hat{\tilde{H}}$ to replace it [CS16]. Suppose that our Hamiltonian can be written as summation of several easy to square root Hamiltonians(we assume that each $\hat{h}_l$ is positive semi definite matrix):

$$\hat{H} = \sum_l \hat{h}_l. \tag{17}$$

We can show that the following effective Hamiltonian behaves as $\sqrt{\hat{H}}$ in the block of $|0\rangle\langle 0|$:

$$\hat{\tilde{H}} = \sum_l \sqrt{\hat{h}_l} \otimes (|0\rangle\langle l| + |l\rangle\langle 0|). \tag{18}$$

Similar to the analysis in Section 3.2, we obtain a complexity of

$$N_{\sigma_H} = \tilde{O}(\beta^2 ||\tilde{H}||^2 polylog(\frac{1}{\epsilon})) \tag{19}$$

3.4 Petz recovery construction with QSVT

[GLM+22] provides a construction of the Petz recovery channel. The original method in [GLM+22] assumes that we have access to copies of the state σ, which in our case is the thermal state σ_H. This will lead to a 'circular argument' if our only goal is to prepare σ_H, but the algorithm is still useful if we wish to follow the dynamics of some initial state under the thermalizing Petz dynamics. In addition, the algorithm described in [GLM+22] does not necessarily requires accessing to copies of σ_H, but only requires access to the block encoding of imaginary time evolution

$$\hat{V}_{\sigma_H} = \begin{bmatrix} e^{-\beta\hat{H}}/r & \cdot \\ \cdot & \cdot \end{bmatrix}. \tag{20}$$

The gate complexity given by [GLM+22] reads

$$\tilde{\mathcal{O}}\left[\sqrt{d_E \kappa_{\mathcal{E}(\sigma_H)}}\left[\kappa_{\mathcal{E}(\sigma_H)} N_{\mathcal{E}(\sigma_H)} + N_{\mathcal{E}} + N_{\sigma_H} \min\left(\kappa_{\sigma_H}, d_E \kappa_{\mathcal{E}(\sigma_H)}/\varepsilon^2\right)\right]\right], \tag{21}$$

where d_E is the Kraus rank of channel d_E is the Kraus rank of channel $\mathcal{E}$, $\kappa_{\mathcal{E}(\sigma)}$ is the inverse of minimum eigen-value of $\mathcal{E}(\sigma_H)$, $N_{\mathcal{E}(\sigma_H)}$ is bounded by $N_{\mathcal{E}} + N_{\sigma_H}$, $N_{\mathcal{E}}$ is the complexity of implementing noise channel $\mathcal{E}$, and N_{σ_H} is the complexity of implementing $\hat{V}_{\sigma_H}$, which is given by Eq. (15) or Eq. (19). So we can simplify the complexity to

$$\tilde{O}[\sqrt{d_E \kappa_{\mathcal{E}(\sigma_H)}}(\kappa_{\mathcal{E}(\sigma_H)} + \kappa_{\sigma_H})[N_{\sigma_H} + N_{\mathcal{E}}]]. \tag{22}$$

wenhao: can be more tight

3.5 Petz recovery construction with QPD

[QR22] provides a alternative way to implement Petz recovery based on a improved quantum polar decomposition (QPD) algorithm. To build the polar decomposition, we first purify the transpose of noise channel as below(which is not unitary in general):

$$\hat{U}_{\mathcal{E}^T} = \sum_i \hat{E}_i^\dagger \otimes |i\rangle_a \langle 0|_a. \tag{23}$$

wenhao: when designing the noise channel, we can go beyond kraus form and design $\hat{U}_{\mathcal{E}^T}$ as a whole

Define the polar decomposition of a operator as

$$\hat{P}(\hat{A}) = \hat{A}(\hat{A}^\dagger \hat{A})^{-\frac{1}{2}} \tag{24}$$

We note that Petz recovery corresponds to $\hat{P}(\hat{U}_{\mathcal{E}^T})$:

$$\mathcal{R}^{\mathcal{E}}_{\sigma_H}(\cdot) = \mathrm{Tr}_a\, \hat{P}(\sqrt{\sigma_H}\hat{U}_{\mathcal{E}^T})(\cdot)\hat{P}(\sqrt{\sigma_H}\hat{U}_{\mathcal{E}^T})^\dagger. \tag{25}$$

[QR22] provides an algorithm to implement the polar decomposition $\hat{P}(\cdot)$ with complexity $O(\frac{\log 1/\epsilon}{\delta})$ wenhao: (definition of δ need to be checked) , so the overall complexity of performing Petz recovery is

$$O(\frac{\log 1/\epsilon}{\delta}(N_{\sigma_H} + N_{\mathcal{E}})) \tag{26}$$

3.6 Converging to steady state

According to fixed point theorem, Eq. (3) implies

$$\lim_{n\to\infty} (\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^n (\rho_0) = \sigma \tag{27}$$

wenhao: here we need to make sure that there is only one steady state (in classical Markov chain, it requires the stochastic process to be ergodic. How about quantum?)

Evaluation of the convergence rate is a non-trivial task both analytically and numerically for the reason that Petz recovery formula contains a term $\mathcal{E}(\sigma_H)^{-\frac{1}{2}}$, which is hard to simplify analytically, and will lead to numerical instability in numerical simulation. Fortunately, we can use some technique to circumvent this and simplify the expression containing only two variables: QEC matrix M and thermal state σ_H.

Mathematical details are delayed to Appendix A. From theorem 1 we know that the convergence of $(\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^n$ is determined by the matrix $\Lambda_{[\mu\mu'],[\nu\nu']}$, which is defined by Eq. (38). Suppose the eigenvalues of $\Lambda_{[\mu\mu'],[\nu\nu']}$ are $\{\lambda_i\}$. Because Λs are Hermitian, the eigenvalues are real. Also, because $(\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^n$ will not diverge, we assume that all eigenvalues satisfy $-1 \le \lambda_i \le 1$.wenhao: (Λ can be further simplified here). We define a gap for this matrix

$$\Delta_\Lambda = \min_{\substack{\lambda \in \text{Spectrum of } \Lambda \\ |\lambda| \neq 1}} 1 - |\lambda|, \tag{28}$$

Then the convergence step of $(\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^n$ is

$$n_c = O(\frac{\ln \epsilon}{\ln(1-\Delta_\Lambda)}) \tag{29}$$

where ϵ is the error between $(\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^{n_c}$ and $\lim_{n\to\infty}(\mathcal{R}^{\mathcal{E}}_{\sigma} \circ \mathcal{E})^n$.

3.7 How to choose the noise channel $\mathcal{E}$

wenhao: TBD: find the best $\mathcal{E}$, maybe numerical methods. wenhao: we should notice that $\hat{U}_{\mathcal{E}^T}$ has another nice property: QEC matrix $M = \hat{U}_{\mathcal{E}^T}\hat{U}^\dagger_{\mathcal{E}^T}$(?). This is another reason we should directly design $\hat{U}_{\mathcal{E}^T}$ rather than Kraus operators.

3.8 Overall complexity

Because we have multiple strategies for each sub-Problem, we show the overall complexity with different combinations in Table 1.

[SL: what is the ϵ dependence in the HS method?]

4 APPLICATION TO LATTICE MODELS

wenhao: TBD: discuss lattice models, especially frustrated Hamiltonians

Table 1: a summary of overall state preparation complexity, for different strategy combinations

Petz recovery construction / imaginary time evolution	QSVT based	QPD based
LCU based	$\tilde{O}(\frac{\sqrt{d_E\kappa_{\mathcal{E}(\sigma_H)}}\left(\kappa_{\mathcal{E}(\sigma_H)}+\kappa_{\sigma_H}\right)}{\ln(1-\Delta_\Lambda)}(\frac{(\beta\|H\|)^2}{\epsilon^3}+N_{\mathcal{E}}))$	$\tilde{O}(\frac{1}{\delta\ln(1-\Delta_\Lambda)}(\frac{(\beta\|H\|)^2}{\epsilon^3}+N_{\mathcal{E}}))$
H-S based	$\tilde{O}(\frac{\sqrt{d_E\kappa_{\mathcal{E}(\sigma_H)}}\left(\kappa_{\mathcal{E}(\sigma_H)}+\kappa_{\sigma_H}\right)}{\ln(1-\Delta_\Lambda)}((\beta\|\tilde{H}\|)^2+N_{\mathcal{E}}))$	$\tilde{O}(\frac{1}{\delta\ln(1-\Delta_\Lambda)}((\beta\|\tilde{H}\|)^2+N_{\mathcal{E}}))$

5 COMPARISON WITH PREVIOUS RESULTS

wenhao: TBD

References

[AAB+24] Yuri Alexeev, Maximilian Amsler, Marco Antonio Barroca, Sanzio Bassini, Torey Battelle, Daan Camps, David Casanova, Young jai Choi, Frederic T Chong, Charles Chung, et al. Quantum-centric supercomputing for materials science: A perspective on challenges and future directions. *Future Generation Computer Systems*, 2024.

[ALL23] Dong An, Jin-Peng Liu, and Lin Lin. Linear combination of hamiltonian simulation for nonunitary dynamics with optimal state preparation cost. *Physical Review Letters*, 131(15):150603, 2023.

[BCL24] Thiago Bergamaschi, Chi-Fang Chen, and Yunchao Liu. Quantum computational advantage with constant-temperature gibbs sampling. *arXiv preprint arXiv:2404.14639*, 2024.

[BCS+20] Dominic W Berry, Andrew M Childs, Yuan Su, Xin Wang, and Nathan Wiebe. Time-dependent hamiltonian simulation with l^1-norm scaling. *Quantum*, 4:254, 2020.

[BLMT24] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. High-temperature gibbs states are unentangled and efficiently preparable. *arXiv preprint arXiv:2403.16850*, 2024.

[CKBG23] Chi-Fang Chen, Michael J Kastoryano, Fernando GSL Brandão, and András Gilyén. Quantum thermal state preparation. *arXiv preprint arXiv:2303.18224*, 2023.

[CKG23] Chi-Fang Chen, Michael J Kastoryano, and András Gilyén. An efficient and exact non-commutative quantum gibbs sampler. *arXiv preprint arXiv:2311.09207*, 2023.

[CS16] Anirban Narayan Chowdhury and Rolando D Somma. Quantum algorithms for gibbs sampling and hitting-time estimation. *arXiv preprint arXiv:1603.02940*, 2016.

[DLL24] Zhiyan Ding, Bowen Li, and Lin Lin. Efficient quantum gibbs samplers with kubo–martin–schwinger detailed balance condition, 2024.

[GLM+22] András Gilyén, Seth Lloyd, Iman Marvian, Yihui Quek, and Mark M. Wilde. Quantum algorithm for petz recovery channels and pretty good measurements. *Physical Review Letters*, 128(22), June 2022.

[MZAW24] Danial Motlagh, Modjtaba Shokrian Zini, Juan Miguel Arrazola, and Nathan Wiebe. Ground state preparation via dynamical cooling. *arXiv preprint arXiv:2404.05810*, 2024.

[QR22] Yihui Quek and Patrick Rebentrost. Fast algorithm for quantum polar decomposition and applications. *Physical Review Research*, 4(1):013144, 2022.

[ZBC23] Daniel Zhang, Jan Lukas Bosse, and Toby Cubitt. Dissipative quantum gibbs sampling. *arXiv preprint arXiv:2304.04526*, 2023.

[ZHLJ24] Guo Zheng, Wenhao He, Gideon Lee, and Liang Jiang. The near-optimal performance of quantum error correction codes, 2024.

A Mathematical Details of Channel $(\mathcal{R}_\sigma^\mathcal{E} \circ \mathcal{E})^n$

We can expect that there are two factors affects the channel $(\mathcal{R}_\sigma^\mathcal{E} \circ \mathcal{E})^n$: one is the choice of σ in Petz recovery, another is the choice of $\mathcal{E}$. We can describe the former one by

$$\sigma_H = \sum_\mu p_\mu |\mu\rangle \langle\mu| \tag{30}$$

, and describe the latter one by QEC matrix, the notion of which is borrowed from quantum error correction. We express the QEC matrix on the same basis as σ_H being diagonalized, i.e. $\{|\mu\rangle\}$:

$$M_{[\mu l],[\nu k]} = \langle\mu| E_l^\dagger E_k |\nu\rangle . \tag{31}$$

We can imagine that given p, M, and basis $\{|\mu\rangle\}$, we can write down channel $(\mathcal{R}_{\sigma_H}^\mathcal{E} \circ \mathcal{E})^n$. In fact, we are able to derive such a expression as the following theorem.

Theorem 1. *Channel* $(\mathcal{R}^{\mathcal{E}}_{\sigma_H} \circ \mathcal{E})^n$ *can be represented as*

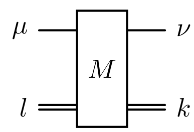
$$(\mathcal{R}^{\mathcal{E}}_{\sigma_H} \circ \mathcal{E})^n(\cdot) = \sum_{\mu_1,\dots,\mu_{n+1}} \sum_{\mu'_1,\dots,\mu'_{n+1}} \Lambda_{[\mu_1\mu'_1],[\mu_2\mu'_2]}\Lambda_{[\mu_2\mu'_2],[\mu_3\mu'_3]}\cdots\Lambda_{[\mu_n\mu'_n],[\mu_{n+1}\mu'_{n+1}]} \tag{32}$$

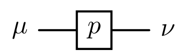
$$p^{\frac{1}{4}}_{\mu_{n+1}} |\mu_{n+1}\rangle \langle\mu_1| p^{-\frac{1}{4}}_{\mu_1}(\cdot)p^{-\frac{1}{4}}_{\mu'_1} |\mu'_1\rangle \langle\mu'_{n+1}| p^{\frac{1}{4}}_{\mu'_{n+1}} \tag{33}$$

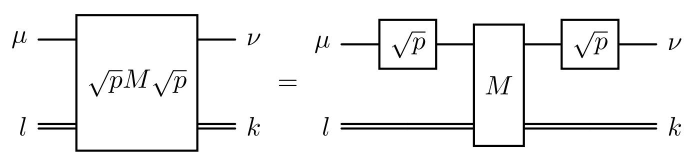
where

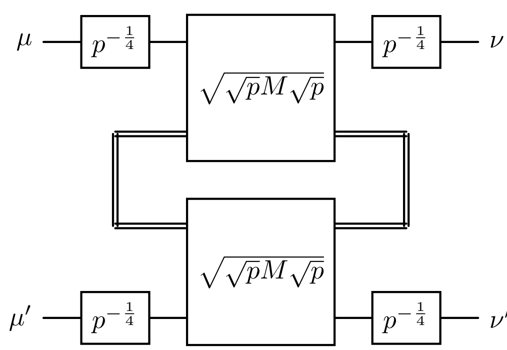
$$\Lambda_{[\mu\mu'],[\nu\nu']} = \sum_{l,k}(p^{-\frac{1}{4}}\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{4}})_{[\mu l],[\nu k]}(p^{-\frac{1}{4}}\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{4}})_{[\mu' l],[\nu' k]} \tag{34}$$

To visulize theorem 1, we can use a tensornetwork representation:

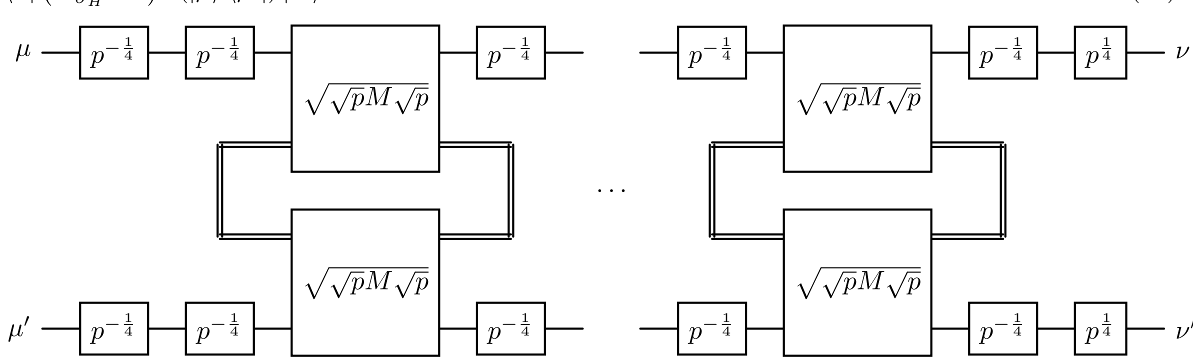
$M \sim M_{[\mu l],[\nu k]} =$  (35)

$p \sim p_\mu \delta_{\mu,\nu} =$  (36)

$\sqrt{p}M\sqrt{p} \sim \sqrt{p_\mu}M_{[\mu l],[\nu k]}\sqrt{p_\nu} =$  (37)

$\Lambda \sim \Lambda_{[\mu\mu'],[\nu\nu']} =$  (38)

$\langle\nu| \left(\mathcal{R}^{\mathcal{E}}_{\sigma_H} \circ \mathcal{E}\right)^n (|\mu\rangle\langle\mu'|) |\nu'\rangle =$ (39)



(40)

We first give two lemmas and prove the theorem in the end.

Lemma 1. *Supposed that σ_H can be diagonalized on basis $\{|\mu\rangle\}$: $\sigma_H = \sum_\mu p_\mu |\mu\rangle\langle\mu|$, and $\{E_i\}$ is Kraus representationof channel $\mathcal{E}$, then we have*

$$\mathcal{E}^{-\frac{1}{2}}(\sigma_H) = \sum_{l_1,l_2,\mu_1,\mu_2} A_{[\mu_1 l_1],[\mu_2,l_2]} E_{l_1} |\mu_1\rangle\langle\mu_2| E_{l_2}^\dagger \tag{41}$$

with matrix

$$A = \sqrt{p}(\sqrt{p}M\sqrt{p})^{-\frac{3}{2}}\sqrt{p}, \tag{42}$$

where p is the matrix of operator σ_H on $\{|\mu\rangle\}$ basis.

Proof. It is sufficient to check Eq.(41) satisfies the definition of inverse square root:

$$\mathcal{E}^{-\frac{1}{2}}(\sigma_H)\mathcal{E}(\sigma_H)\mathcal{E}^{-\frac{1}{2}}(\sigma_H) = P_{\mathcal{E}(\sigma_H)}, \tag{43}$$

where the RHS can be expanded on $E_i|\mu\rangle$ basis as following

$$RHS = \sum_{\mu_1,l_1,\mu_2,l_2} (M^{-1})_{[\mu_1 l_1],[\mu_2,l_2]} E_{l_1}|\mu_1\rangle\langle\mu_2| E_{l_2}^\dagger, \tag{44}$$

while the LHS can be simplified as

$$LHS = \sum_{l_1,l_2,\mu_1,\mu_2} A_{[\mu_1 l_1],[\mu_2,l_2]} E_{l_1}|\mu_1\rangle\langle\mu_2| E_{l_2}^\dagger \sum_{\mu,i} p_\mu E_i|\mu\rangle\langle\mu| E_i^\dagger \sum_{l_1',l_2',\mu_1',\mu_2'} A_{[\mu_1' l_1'],[\mu_2',l_2']} E_{l_1'}|\mu_1'\rangle\langle\mu_2'| E_{l_2'}^\dagger \tag{45}$$

$$= \sum_{\mu_1,l_1,\mu_2,l_2} (AMpMA)_{[\mu_1 l_1],[\mu_2,l_2]} E_{l_1}|\mu_1\rangle\langle\mu_2| E_{l_2}^\dagger \tag{46}$$

$$= RHS \tag{47}$$

□

Lemma 2. *The Kraus representation of composition map $\mathcal{R}_{\sigma_H}^{\mathcal{E}} \circ \mathcal{E}$ reads*

$$\{\sum_{\mu,\nu} Q_{[\mu l],[\nu k]} |\mu\rangle\langle\nu|\}_{l,k} \tag{48}$$

where

$$Q_{[\mu l],[\nu k]} = (\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{2}})_{[\mu l],[\nu k]} \tag{49}$$

. Here we assume that input state of $\mathcal{R}_{\sigma_H}^{\mathcal{E}} \circ \mathcal{E}$ is on the support of basis $\{|\mu\rangle\}$.

Proof. Kraus operator of of a composition map can be written as the product of each Kraus oeprator. So the Kraus representation of composition map $\mathcal{R}_{\sigma_H}^{\mathcal{E}} \circ \mathcal{E}$ reads

$$\sigma_H{}^{\frac{1}{2}} E_l^\dagger \mathcal{E}(\sigma_H)^{-\frac{1}{2}} E_k = \sum_\mu \sqrt{p_\mu}|\mu\rangle\langle\mu| E_l^\dagger \sum_{l_1,l_2} A_{[\mu_1 l_1],[\mu_2,l_2]} E_{l_1}|\mu_1\rangle\langle\mu_2| E_{l_2}^\dagger E_k \sum_\nu |\nu\rangle\langle\nu| \tag{50}$$

$$= \sum_{\mu,\nu} Q_{[\mu l],[\nu k]} |\mu\rangle\langle\nu|, \tag{51}$$

where the matrix

$$Q_{[\mu l],[\nu k]} = \sum_{l_1,l_2} \sqrt{p_\mu} A_{[\mu_1 l_1],[\mu_2,l_2]} M_{[\mu l],[\mu_1 l_1]} M_{[\mu_2 l_2],[\nu k]} \tag{52}$$

$$= (\sqrt{p}MAM)_{[\mu l],[\nu k]} \tag{53}$$

$$= (\sqrt{p}M\sqrt{p}(\sqrt{p}M\sqrt{p})^{-\frac{3}{2}}\sqrt{p}M)_{[\mu l],[\nu k]} \tag{54}$$

$$= (\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{2}})_{[\mu l],[\nu k]} \tag{55}$$

□

Then we can give the proof of Theorem 1

Proof. Using Lemma 2, we have

$$(\mathcal{R}_{\sigma_H}^{\mathcal{E}} \circ \mathcal{E})^n(\cdot) = \sum_{\mu_1,\ldots,\mu_{n+1}} \sum_{\mu'_1,\ldots,\mu'_{n+1}} \sum_{l_1,\ldots l_n} \sum_{k_1,\ldots,k_n} \tag{56}$$
$$(Q_{[\mu_{n+1}l_n],[\mu_n k_n]} \cdots Q_{[\mu_3 l_2],[\mu_2 k_2]} Q_{[\mu_2 l_1],[\mu_1 k_1]}) \tag{57}$$
$$|\mu_{n+1}\rangle \langle \mu_1| (\cdot) |\mu'_1\rangle \langle \mu'_{n+1}| \tag{58}$$
$$(Q^\dagger_{[\mu'_1 k_1],[\mu'_2 l_1]} Q^\dagger_{[\mu'_2 k_2],[\mu'_3 l_2]} \cdots Q^\dagger_{[\mu'_n k_n],[\mu'_{n+1} l_n]}), \tag{59}$$

where

$$Q_{[\mu l],[\nu k]} = (\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{2}})_{[\mu l],[\nu k]} \tag{60}$$
$$Q^\dagger_{[\mu l],[\nu k]} = (p^{-\frac{1}{2}}\sqrt{\sqrt{p}M\sqrt{p}})_{[\mu l],[\nu k]}. \tag{61}$$

In order to symmetrize Q, we can try to split every $p^{-\frac{1}{2}}$ into two $p^{-\frac{1}{4}}$s, and define

$$\tilde{Q}_{[\mu l],[\nu k]} = (p^{-\frac{1}{4}}\sqrt{\sqrt{p}M\sqrt{p}}p^{-\frac{1}{4}})_{[\mu l],[\nu k]}, \tag{62}$$

with which we can rewrite Eq. (56) as

$$(\mathcal{R}_{\sigma_H}^{\mathcal{E}} \circ \mathcal{E})^n(\cdot) = \sum_{\mu_1,\ldots,\mu_{n+1}} \sum_{\mu'_1,\ldots,\mu'_{n+1}} \sum_{l_1,\ldots l_n} \sum_{k_1,\ldots,k_n} \tag{63}$$
$$(\tilde{Q}_{[\mu_{n+1}l_n],[\mu_n k_n]} \cdots \tilde{Q}_{[\mu_3 l_2],[\mu_2 k_2]} \tilde{Q}_{[\mu_2 l_1],[\mu_1 k_1]}) \tag{64}$$
$$p^{\frac{1}{4}}_{\mu_{n+1}} |\mu_{n+1}\rangle \langle \mu_1| p^{-\frac{1}{4}}_{\mu_1} (\cdot) p^{-\frac{1}{4}}_{\mu'_1} |\mu'_1\rangle \langle \mu'_{n+1}| p^{\frac{1}{4}}_{\mu'_{n+1}} \tag{65}$$
$$(\tilde{Q}_{[\mu'_1 k_1],[\mu'_2 l_1]} \tilde{Q}_{[\mu'_2 k_2],[\mu'_3 l_2]} \cdots \tilde{Q}_{[\mu'_n k_n],[\mu'_{n+1} l_n]}). \tag{66}$$

We can furthur simplify it by grouping some indeces and thus conclude the proof. □