Reply to Report 1 by Referee 1:

Report: *In the first half of this paper, which is mostly a review, the authors discuss four different types of entanglement entropy. Several examples are worked out in the second half of the paper. What I miss in this paper is the physical motivation for introducing these entropies. Why do we need them? The lack of motivation makes reading the first half confusing. The worked out examples are in the second half of the paper, but it is not clear exactly what is calculated. In order to reproduce the figures, I have to guess what the authors are doing, and it should not be like that. The authors should be very clear about the equation they have been using to calculate the entropies.*

Authors: We are very grateful to the referee for the critical review and constructive assessment of our manuscript. We have carefully revised the manuscript in response to the referee's comments. We clarify the motivation for introducing these entropies of the transition matrix, namely to quantify how much information can be transferred through such post-selected processes. We also provide detailed procedures and equations used in the computations in the second half of the manuscript. All changes in our manuscript have been highlighted in red. Our point-by-point responses to the referee's questions are given below.

1. **Q:** *It would be a good idea to shorten the first half of the paper. What should be clear are the general properties that a transition matrix should satisfies. From reading the paper I get the idea that it can be anything. The connection the authors make with teleportation is confusing because the examples are not at all related to teleportation.*

   A: We thank the referee for your suggestion. We have shortened the Introduction greatly to make the motivation more clear and to present the content in the main text more concise. We have also removed some redundant discussions in Secs. 2.1 and 2.2.

   In Sec. 2.1, we now emphasize that τ depends strictly on the bipartite states $|\psi_2\rangle_{ab}$ and $|\psi_1\rangle_{bc}$. It is not an independently defined operator, but rather an effective object induced by the post-selection on subsystem b. We also clarify that the transition matrix τ is not necessarily a square matrix. Even for square cases, τ is generally non-Hermitian except for some special situations in which $|\tilde{\gamma}_i\rangle = \pm\,|\tilde{\alpha}_i\rangle$ holds for all the indices i with $z_i > 0$. As a result, τ is not an independent random matrix, as further addressed on Page 18.

   In Sec. 2.2, we point out that for non-Hermitian τ, the spectrum may include negative and complex values. As a result, the trace $\mathrm{Tr}\,\tau$ is not unity in general, but can be negative, complex, arbitrarily small or even vanish, leading to the (modified) pseudo entropy becoming negative, complex or divergent.

   We also removed the detailed review of quantum teleportation. Quantum teleportation is a quantum channel that faithfully transfers a quantum state, while the post-selection process described by a transition matrix is completely positive but

not trace-preserving. It therefore does not define a quantum channel and transfers only partial information. Since the subsequent examples are not directly related to teleportation, this comparison has been removed to avoid potential confusion.

2. **Q:** *Eq. (6): $S[\rho||\sigma]$ is not defined. Please give an explicit definition.*
A: Below Eq.(6), we add the definition of the relative entropy by

$$S[\rho||\sigma] = \mathrm{Tr}\left[\rho \ln \rho\right] - \mathrm{Tr}\left[\rho \ln \sigma\right].$$

Above Eq. (9), we also add that for when σ is a maximally mixed state, then any unitary operation leaves the relative entropy unchanged.

3. **Q:** *Line above Eq. (19): we do not need "novel".*
A: We agree with the referee. We have removed the "novel" above Eq. (18). $\bar{\tau}$ constructed via singular value decomposition is proposed in ref. [14].

4. **Q:** *It is not clear what the authors mean by distillation and why it is useful. Please explain and motivate.*
A: On Page 3, we have made it clear that the distillation of transition matrices is an analogy to the entanglement distillation of pure states. In Sec. 3.1, we have reviewed that the entanglement distillation refers to the process of transforming the m-copy entangled states into some amount of EPR pairs through LOCC. In the asymptotic limit $m \to \infty$, the entanglement entropy S in base 2 serves as the asymptotic rate, quantifying the average number of EPR pairs that can be extracted from each copy of the state.
By analogy, the distillation of transition matrices refers to a procedure in which many copies of transition matrices are asymptotically concentrated into a lower-dimensional transition matrix. This process is interpreted as being equivalent to distilling a certain number of two-dimensional transition matrices, each of which has the corresponding entropy $\ln 2$. The total entropy of many copies of transition matrices is thereby asymptotically concentrated into that of the lower-dimensional transition matrix. If some entropy measures of the transition matrix we choose can take the negative, complex or even divergent values, as for example, the pseudo and modified pseudo entropies, then distillation of transition matrices does not apply to such entropy measures. Some entropy measures like SVD and ABB entropies, admit such a distillation interpretation, meaning that the corresponding entropy of a single transition matrix is real, non-negative, and bounded by the dimension of the Hilbert space.
About usefulness, distillation offers a clear and effective description of complicated transition matrices and also reveals the irreversibility of additional non-unitary operations (Sec. 2.3). That's why distillation of transition matrices is useful.
The entanglement distillation of pure states is physically operational. By generalized measurement via LOCC, many copies of a pure state will collapse into a maximally entangled state with the highest probability in the sense of quantum mechanics. However, the so-called probability of concentration appearing in the distillation for SVD entropy is not the genuine probability of measurement in the framework of quantum mechanics. This issue motivates us to consider another

entropy measure of transition matrix, i.e., ABB entropy, for which the distillation of transition matrices admits a clear probability interpretation.

5. **Q:** *Explain Eq. (36). I assume that $d_{1\mathbf{k}}$ is still given by the binomial factor. Mention at least Stirling. What is exactly $|0\rangle$ in (54). Can the author add an appendix with an expression for (54) with all indices written out explicitly including the tensor products.*
A: We thank the referee for the detailed and helpful comments. We address each point as follows.
(1). Explanation of Eq. (37)(Eq. (36) in the previous version):
We have added the explicit asymptotic form of Stirling's approximation used here above Eq.(34), namely,

$$m! \sim \sqrt{2\pi m}(m/e)^m, \quad k_i! \sim \sqrt{2\pi k_i}(k_i/e)^{k_i},$$

valid for large m and k_i. In addition, we have also corrected the typos in Eq.(34) and (35). With the multinominal factor in Eq.(32),

$$d_{1\mathbf{k}} = \frac{m!}{k_1!k_2!\cdots k_{d_1}!},$$

and $\mathbf{k}^* = \left(mx_1^2, \cdots, mx_{d_1}^2\right)$, the detailed derivation of Eq.(37) is written as

$$\begin{aligned}\lim_{m\to\infty} \frac{\ln d_{1\mathbf{k}^*}}{m} &= \frac{1}{m}\ln\frac{\sqrt{2\pi m}}{\prod_i^{d_i}\sqrt{2\pi k_i^*}} + \ln m - \sum_{i=1}^{d_i}\frac{k_i^*\ln k_i^*}{m} \\ &= -\sum_{i=1}^{d_1} x_i^2 \ln x_i^2 = S_1,\end{aligned}$$

where the first term can be ignored in the large-m limit and we used $\sum_i^{d_i} x_i^2 = 1$ at the second equality. Below Eq. (37), we have also added the explanation that $\log_2 d_{1\mathbf{k}^*}$ measures the expected number of 2-dimensional EPR pairs that could be distilled from $|\psi_1\rangle_{bc}^{\otimes m}$ and the multinomial factor Eq. (32) for $\mathbf{k}_*$ is used to compute the limit Eq. (37).
(2). Meaning of the reference state in Eq. (54):
In Eq. (55) (Eq. (54) in the previous version), the specific choice of the reference state $|0\rangle$ is not important. The Haar random state in this paper is defined by applying a unitary matrix U, drawn from the circular unitary ensemble (CUE) $\mathcal{E}$, to an arbitrary reference state $|0\rangle$. Let's choose another reference state $|0'\rangle = U'|0\rangle$, which is related to $|0\rangle$ by an unitary transformation U'. But the ensemble of Haar random states is invariant

$$\{U|0'\rangle \mid U \in \mathcal{E}\} = \{UU'|0\rangle \mid U \in \mathcal{E}\} = \{U|0\rangle \mid U \in \mathcal{E}\}$$

since CUE is invariant under left multiplication.

6. **Q:** *Eq. (55): dU and dV do not have indices but U and V do. Are they the same?*
A: Yes, they are the same. In Eq.(56) (Eq.(55) in the previous version), dU and dV correspond to U_{bc} and V_{ab} appearing in the integrand. To avoid ambiguity, we have explicitly restored the labels and now write them as dU_{bc} and dV_{ab}. By the way, a and b are labels denoting two subsystems, rather than indices of a matrix. Moreover, since U and V are independently drawn from their respective ensembles, the ensemble average over U_{bc} is independent of that over V_{ab}.

7. **Q:** *Fig. 5: Why some curves are symmetric and other are not. Please explain exactly which equations have been used for these figures.*
A: Thanks for referee's suggestion. In the caption of Fig. 5, we have added the statement that we are considering the transition matrix constructed from two independent Haar random states in Eq. (55),

$$\tau = \mathrm{Tr}_b \left[U_{ba} \left|0\right\rangle_{ba} \left\langle 0\right|_{ab} V_{ab}^\dagger \right].$$

To compute the four entropies, we substituted τ into Eqs. (14)(16)(19)(10) respectively.
We now explain why the ensemble-averaged (modified) pseudo entropy curves are symmetric under the exchange $d_1 \leftrightarrow d_2$, whereas the ensemble-averaged SVD and ABB entropy curves are not. Writing the bipartite states $\left|\psi_1\right\rangle_{ab}$ and $\left|\psi_2\right\rangle_{ab}$ with subsystem dimensions (d_1, d_2) in matrix form as $\psi \in \mathbb{C}^{d_1 \times d_2}$ and $\phi \in \mathbb{C}^{d_1 \times d_2}$, tracing out b yields

$$\tau_a = \psi\phi^\dagger \in \mathbb{C}^{d_1 \times d_1},$$

while tracing out a gives

$$\tau_b = \left(\phi^\dagger\psi\right)^T \in \mathbb{C}^{d_2 \times d_2}.$$

By the Sylvester determinant theorem, it can be verified that τ_a and τ_b share the same set of nonzero eigenvalues. Consequently, the pseudo entropy and modified pseudo entropy, both of which depend only on the spectrum of τ, are invariant under exchanging the subsystem dimensions. After ensemble averaging, this leads to

$$\overline{S_{\mathrm{P}}\left[\tau_a(d_1;d_2)\right]} = \overline{S_{\mathrm{P}}\left[\tau_a(d_2;d_1)\right]}, \quad \overline{S_{\mathrm{MP}}\left[\tau_a(d_1;d_2)\right]} = \overline{S_{\mathrm{MP}}\left[\tau_a(d_2;d_1)\right]}.$$

In contrast, the SVD and ABB entropies depend on the spectra of $\tau\tau^\dagger$ and related normalized operators. In general, $\tau_a\tau_a^\dagger$ and $\tau_b\tau_b^\dagger$ do not share the same nonzero spectra, leading to that $\bar{\tau}_a$ and $\bar{\tau}_b$, as well as $\tilde{\tau}_a\tilde{\tau}_a^\dagger$ and $\tilde{\tau}_b\tilde{\tau}_b^\dagger$ generally do not have the same spectra. As a result,

$$\overline{S_{\mathrm{ABB}}\left[\tau_a(d_1;d_2)\right]} \neq \overline{S_{\mathrm{ABB}}\left[\tau_a(d_2;d_1)\right]}, \quad \overline{S_{\mathrm{SVD}}\left[\tau_a(d_1;d_2)\right]} \neq \overline{S_{\mathrm{SVD}}\left[\tau_a(d_2;d_1)\right]},$$

which explains the asymmetry of the ensemble-averaged ABB and SVD entropy observed in Fig. 5. The detailed proof can be found in App. C. In addition, we have included the explanations of why the (modified) pseudo entropy are symmetric on Page 20 and why the SVD and ABB entropies are asymmetric on Page 21.

8. **Q:** *What is the transition matrix for which the eigenvalues in Fig. 6 have been calculated ?*
   A: The eigenvalues shown in Fig. 6 are calculated from the transition matrix constructed from two independent Haar random states in Eq. (55). Specifically, for each disorder realization, we calculated the eigenvalues $\{\lambda_i\}$ of $\hat{\tau}$, the corresponding quantities $\left\{\left|\lambda_i - \bar{\lambda}\right|\right\}$, the eigenvalues $\{q_i\}$ of $\bar{\tau}$, and the eigenvalues $\{p_i\}$ of $\tilde{\tau}\tilde{\tau}^\dagger$. This information has now been stated explicitly in the figure caption.

9. **Q:** *a and b are the left and right eigenvectors of the Ginibre ensemble according to (64) and must have the same dimension, but this does not square with Fig. 7. Please explain exactly what is calculated and which equations have been used. Also explain why some curves appear symmetric and others do not.*
   A: Here, a and b refer to two subsystems with different dimensions in general. They are not left and right eigenvectors. $|L_n\rangle$ and $|R_n\rangle$ represent left and right eigenvectors. Below Eq. (65), we stress that since $|\psi_1\rangle = |L_n\rangle$ and $|\psi_2\rangle = |R_n\rangle$ have the same dimension, the transition matrix τ is square, but subsystem a and b can have different dimensions.
   To clarify which transition matrix is used for Fig. 7, we add a new Eq. (66):

$$(H - w)\,|R_{n_w}\rangle = (\mathcal{E}_{n_w} - w)\,|R_{n_w}\rangle\,, \quad \langle L_{n_w}|\,(H - w) = (\mathcal{E}_{n_w} - w)\,\langle L_{n_w}|\,,$$

   and Eq. (67):

$$\tau_w = \mathrm{Tr}_b\ |R_{n_w}\rangle\,\langle L_{n_w}|\,.$$

   For Fig. 7, we calculate the four entropies of τ_w in Eq. (67), and then perform ensemble averages over the Ginibre ensemble. In the numerical calculation, given a specific complex number w with $|w| < 1$, we single out the eigenvalue $\mathcal{E}_{n_w}$ of H, satisfying that $|\mathcal{E}_{n_w} - w|$ is the smallest among all the eigenvalues of H, and then obtain the left and right eigenstates $|L_{n_w}\rangle$ and $|R_{n_w}\rangle$ corresponding to $\mathcal{E}_{n_w}$. We have also added the explanation in the caption of Fig. 7.
   The reason of why some curves are symmetric while others are not is the same as explained in our previous reply to Question 7, and we have also added a statement on the asymmtry of the SVD and ABB entropies and symmetry of the (modified)pseudo entropy on Page 24.

10. **Q:** *I have the same questions for Fig. 7.*
    A: We suppose you refer to Fig. 8. The answer is the same as the last one.

11. **Q:** *3 Lines below Fig. 7 (6 lines below the title of Sec. 5.1 for the revised manuscript): It is not class A of the AZ classification but class A of the Bernard-LeClair classiufication of non hermitian Random matrices which was completed by Ueda and collaborators.*
    A: We thank the referee for pointing out this mistake. We have corrected the text to refer to class A in Bernard-LeClair classification of non-Hermitian random matrices, rather than the Altland–Zirnbauer (AZ) classification. In addition, we have added the following two references to clarify the classification scheme and

its later completion by Ueda and collaborators:
[104]: D. Bernard and A. LeClair, A classification of non-hermitian random matrices, in Statistical Field Theories, pp. 207–214, Springer (2002).
[105]: K. Kawabata, K. Shiozaki, M. Ueda and M. Sato, Symmetry and Topology in Non-Hermitian Physics, Phys. Rev. X 9 (2019) 041015 [1812.09133].

12. **Q:** *What is the transition matrix used for Fig. 8? Give explicit equations.*
A: For Fig. 8, the transition matrix is still $\tau_w = \mathrm{Tr}_b\, |R_{n_w}\rangle \langle L_{n_w}|$ in Eq. (67), which we have clarified in its caption.

13. **Q:** *4 lines below (67) : change AZ class to Bernard-LeClair class.*
A: We thank the referee for pointing out this. We have also corrected the text four lines below Eq.(70)(Eq.(67) in the previous version) to refer to the Bernard-LeClair classification instead of AZ classification.

14. **Q:** *Could the authors explain exactly which equations are used to obtain Figs. 9-10? It should be done in such a way that the figures can be reproduced without any guess work.*
A: We thank the referee for this important suggestion regarding reproducibility. To clarify which equations are used to obtain Fig. 9 and 10, we have added an explicit statement in the fourth paragraph of Sec. 5.2, specifying that the transition matrix used in these figures is the transition matrix τ_w according to Eq. (67). This matrix is constructed from the bi-orthogonal eigenstates $|R_{n_w}\rangle$ and $\langle L_{n_w}|$ of the non-Hermitian SYK(nSYK) model, whose eigenvalue is closest to w.

15. **Q:** *It is not clear which eqs. are used for Fig. 10. Please provided all details for the transition operator and how it is obtained from the nSYK.*
A: The answer is the same as the last one. The transition matrix is τ_w, given in Eq. (67), $\tau_w = \mathrm{Tr}_b\, |R_{n_w}\rangle \langle L_{n_w}|$. The bi-orthogonal eigenstates are obtained by the same method as in Eq. (66) of Sec. 5.1. We have also added the explanation in the captions of Fig. 9 and 10.

16. **Q:** *For the curves in Figs.14 and 15, do the authors only consider two states or more than two? The exceptional point probably depends on the realization. Is there only one realization? Please state explicitly in the caption.*
A: The answers to both questions are yes. We only consider two branches of states in Fig. 14 and 15. The locations of exceptional points indeed depend on disorder realizations. In each realization, different pairs of branches also collide at different locations of exceptional points. We have mentioned that in the captions of Figs. 13 and 14.
We address that different disorder realizations lead to different spectra, and different pairs of branches undergo the transition at different exceptional points. So we are unable to perform ensemble averaging in sec. 6.2, subjected to the fluctuation of the location of the exceptional points and the choices of branches. But we find that the qualitative nature of the transition, as well as the entropies are universal. We have mentioned this point on page 30.