

Erasing Quantum Coherence: An Operational Approach

Uttam Singh,^{1,*} Manabendra Nath Bera,^{2,1} Avijit Misra,¹ and Arun Kumar Pati¹

¹*Harish-Chandra Research Institute, Allahabad, 211019, India*

²*ICFO-The Institute of Photonic Science, 08860 Castelldefels (Barcelona), Spain*

Despite being one of the hallmarks of quantum physics, there is a lack of operational interpretations of quantum coherence. Here we provide an operational interpretation of coherence of a quantum system, in terms of the amount of noise that is to be injected in order to fully decohere it. In particular, we show that in the asymptotic limit, the minimum amount of noise that is required to fully decohere a quantum system, is equal to the relative entropy of coherence. This quantifies the erasure cost of quantum coherence. We employ the entropy exchanged between system and environment during the decohering operation and the memory required to store the information about the decohering operation as the quantifiers of noise. We show that both the quantifiers yield the same cost of erasing coherence in the asymptotic limit. The relative entropy of coherence, hence, is endowed with a thermodynamical and operational interpretation.

I. INTRODUCTION

With our ever increasing abilities to control systems at smaller and smaller scales, the quantum properties like quantum coherence and quantum entanglement make their presence felt more and more prominently. Recent developments in thermodynamics of nano scale systems suggest that the quantum coherence plays an essential role in determining the state transformations of these systems and more importantly, in providing a family of second laws of thermodynamics [1–10]. Also, the phenomenon of quantum coherence has been arguably attributed to the efficient functioning of some complex biological systems [11–16]. Given the importance of quantum coherence, a formal structure of coherence resource theory is developed in recent years [17–28]. There are two inequivalent frameworks to characterize quantum coherence. The first framework is based on a set of incoherent operations as free operations and a set of freely available incoherent states [19]. This formalism has been successfully applied in the context of quantum entanglement, providing further a family of coherence monotones based on entanglement monotones [24] and quantification of the wave-particle duality [29]. Moreover, a class of maximally coherent mixed states is found, for which coherence and mixedness satisfy a complementarity relation, following this resource theory of coherence [30]. The second formalism is based on the resource theory of asymmetry [17, 18, 22], where operations are restricted to phase insensitive operations and symmetric states are free resources [18]. This formalism has been successfully used in the quantum thermodynamics [6, 7].

But there is no prevalent consensus to which of the two formalisms are better suited, in general, to understand most of the phenomena where coherence plays an important role. The reason for this lack of general consensus can be attributed to the hitherto paucity of operational interpretations of quantum coherence. In the quantum information theory, to equip a particular “resource” of interest with an operational meaning, consideration of thermodynamic cost of destroying (erasing) the “resource”, turns out to be very fruitful and far reaching

[31–37]. For example, the Landauer erasure principle [31] has been a central one in laying the foundation of physics of information theory. Similarly, an operational definition of total correlation, classical correlation and quantum correlation is obtained independently in Refs. [38] and [39], via consideration of thermodynamic cost to erase these correlations. In addition, the thermodynamic cost of erasing quantum correlation needs entropy production has been shown in Ref. [40]. This approach has also been successfully applied to private quantum decoupling [41] and recently to markovianization [42]. These tasks inevitably use randomization of quantum systems [43–45].

In this work, we provide an operational interpretation of quantum coherence in terms of the amount of noise that has to be dumped in to the system such that it decoheres completely. We consider two different measures to quantify the amount of noise in the process of decohering a quantum system: the entropy exchange between system and environment during the decohering operation [46, 47] and the memory required to store the information about the decohering operation [38]. We show that in the asymptotic limit, both these measures yields the same minimal cost of erasing coherence (the minimal noise required to fully decohere the system) and the minimum cost is given by the relative entropy of coherence [19]. Thus, the relative entropy of coherence of a quantum system can be interpreted as the minimal (work) cost of decohering the quantum system completely. As a consequence, the restrictions imposed in the resource theory of coherence [19], namely the allowed operations being incoherent operations and free states being incoherent states, yield a measure of coherence that has an operational significance.

At this point it is worth noting that the erasure of information (in form of correlations or coherence) has connection to the no-hiding theorem [48, 49] that applies to any process hiding a quantum state, whether by randomization, thermalization or any other procedure. The no-hiding theorem [48, 49] that generalises Landauer’s principle [31] offers insight into the nature of thermalization processes in comparison with insights provided by Landauer’s principle in the resolution of Maxwell’s demon. Also, the thermalization process is a particular kind of decohering process in energy eigenbasis and hence our results may have deep connections with the no-

* uttamsingh@hri.res.in

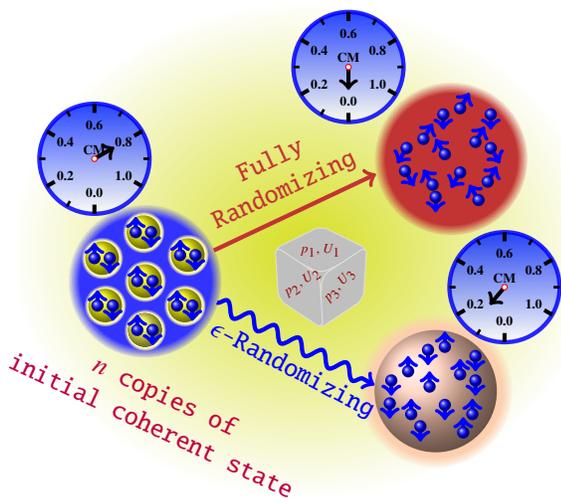


FIG. 1. (Color online) The schematic of the fully and ϵ -randomization operations. The figure depicts that if we start with n copies of any state (coherent or incoherent) and pass them through some randomizing map, then the n copies decohere completely if the map is fully randomizing and if the map is ϵ -decohering then the n copies come very close to the fully decohered state keeping some amount of coherence which is close to zero. We show that in both the cases the minimum amount of noise that is required is same and is equal to relative entropy of coherence in asymptotic limit.

hiding theorem.

The paper is organized as follows: In section II, we give a brief outline of the concepts required to understand the process of erasure of quantum coherence. In section III, we present our main result of obtaining minimal (work) cost of erasing coherence of a quantum system or of decohering a quantum system completely. In section IV, we describe the connection of erasing coherence to thermodynamics and the Landauer erasure principle. Finally we conclude in section V with overview and implications of the results presented in the paper.

II. PRELIMINARIES: VARIOUS DEFINITIONS

Here, we briefly give an account of the concepts that are required to derive our main results. In this paper we will be concerned with the resource theory of coherence as in Ref. [19].

Quantum coherence:— The theory of quantum coherence, starts with fixing a reference basis as coherence is inherently a basis dependent quantity [19]. For a given reference basis $\{|a\rangle\}$, the set of incoherent states \mathcal{I} is defined as the set of all the states of the form $\rho_{\mathcal{I}} = \sum_a p_a |a\rangle\langle a|$, where $\{p_a\}$ is a probability distribution, i.e., $p_a \geq 0$, $\sum_a p_a = 1$, and the incoherent operations $\Lambda^{\mathcal{I}}$ are defined as completely positive trace preserving (CPTP) maps that map the set of incoherent states onto itself. The bonafide measures of coherence that emerge from this theory include the l_1 norm and the relative entropy of coherence [19]. The relative entropy of coherence of any

state ρ is given by [19]

$$C_r(\rho) = H(\rho^d) - H(\rho), \quad (1)$$

where $H(\rho) = -\text{Tr}(\rho \ln \rho)$, is the von Neumann entropy and $\rho^d = \sum_a \langle a|\rho|a\rangle\langle a|$ is the diagonal part of ρ in the reference basis $\{|a\rangle\}$. Furthermore, the maximally coherent state is defined by $|\psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle$, for which $C_r(|\psi_d\rangle\langle\psi_d|) = \ln d$ and the class of maximally coherent mixed states (MCMS) is given by $\rho_p := (1-p)\mathbb{I}_{d \times d}/d + p|\psi_d\rangle\langle\psi_d|$, $0 \leq p \leq 1$, for which $C_r(\rho_p) = \ln d - S(\rho_p)$ [30].

Before we proceed further to present our main result, we would like to give an illustration of the process of fully decohering a qubit quantum system. Let us first fix the reference basis to be the computational basis, $\{|a\rangle\}$ ($a = 0, 1$). Now, consider a qubit system in the state

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} \sum_{a=0}^1 |a\rangle, \quad (2)$$

which is a maximally coherent state in two dimensions and relative entropy of coherence of this state is one bit. Suppose that we want to fully decohere this state, i.e., to erase the coherence of this state. This can be achieved by applying two incoherent unitary transformations \mathbb{I}_2 and σ_z with equal probability. After this action the state becomes

$$\rho = \frac{1}{2} |\psi_2\rangle\langle\psi_2| + \frac{1}{2} \sigma_z |\psi_2\rangle\langle\psi_2| \sigma_z = \frac{1}{2} \mathbb{I}_2. \quad (3)$$

The relative entropy of coherence of this state is zero. This means that applying two incoherent unitary operations probabilistically, with equal probability, will suffice to erase the coherence of the state, given in Eq. (2) (see Fig. 1). Similarly, for MCMS in two dimensions [30], applying two incoherent unitary transformations \mathbb{I}_2 and σ_z with equal probability, yields $\mathbb{I}_2/2$. Also, it can be seen that for a d dimensional quantum system, an ensemble of unitary transformations exists that can decohere any state ρ of the system completely. Now, define the operators \hat{X} and \hat{Z} in the fixed basis, say $\{|1\rangle, \dots, |d\rangle\}$, via

$$\hat{X}|j\rangle = |j \oplus 1\rangle, \text{ and } \hat{Z}|j\rangle = e^{\frac{2\pi i j}{d}} |j\rangle, \quad (4)$$

where \oplus denotes addition modulo d . It is known that [45]

$$\rho \rightarrow \frac{1}{d^2} \sum_{j=1}^d \sum_{k=1}^d \hat{X}^k \hat{Z}^j \rho \hat{Z}^{j\dagger} \hat{X}^{k\dagger} = \frac{1}{d} \mathbb{I}_d. \quad (5)$$

Thus, the incoherent operation $\{\frac{1}{d^2}, \hat{X}^k \hat{Z}^j\}_{jk}$ perfectly randomizes any state (see Fig. 1). However, the no-hiding theorem [48] tells that the original state can be found in the ancilla Hilbert space upto local unitary operation and in fact, this has been experimentally tested [49]. Therefore, the coherence of the original state which is apparently lost, can be found in the ancilla. To completely erase the coherence one has to dump the ancilla subsystem where the original information resides. That will involve thermodynamic cost.

But, what is the cost to be paid in order to implement this probabilistic incoherent operation or how much noise does this operation inject into the system? One possibility, is to consider the amount of information needed to implement this (erasing) operation, which is equal to the Shannon entropy, $H(p = 1/2) = 1$ bit for the qubit example that we have considered above. Therefore, one can say that applying a probabilistic operation consisting of two elements, with equal probability, costs one bit of information or injects one bit of noise in the system. Similarly, for a qudit system, we can achieve exact randomization via a map of the form Eq. (5). The entropy that this map injects in the system as quantified by the amount of information needed to implement it, is given by $H(p = 1/d^2) = 2 \log_2 d$ bits. Clearly, the state independent randomization over estimates the amount of noise that is required to decohere the state (cf. qubit and qudit case). Also, this cost is independent of the nature of the operation, i.e., whether the operation is incoherent, unitary etc. The other choice to quantify the amount of noise injected in the system can be obtained based on exchange entropy as in Refs. [38, 46, 47]. As we show below, the exchange entropy is smaller than $H(p)$.

Exchange entropy:– The exchange entropy [46, 47] is defined as the amount of entropy that any CPTP map R injects into the system S which passes through the channel R . To introduce this measure, we first purify the system state ρ^S by considering a reference system Z such that $\rho^S = \text{Tr}_Z |\psi\rangle \langle \psi|^{SZ}$. Now the entropy that the map R injects into the system is defined as

$$H_e(R, \rho^S) := H\left((R \otimes \mathbb{I}^Z)[|\psi\rangle^{SZ}]\right), \quad (6)$$

where \mathbb{I}_Z is the identity matrix on the reference system Z and H is the von Neumann entropy. The exchange entropy has been successfully employed in gaining insights in security of cryptographic protocols [46, 47], in determining cost of erasing total, classical and quantum correlations [38] and in connection to concurrence [50]. Let R be comprised of random unitary ensemble $\{p_i, u_i\}_{i=1}^N$. The exchange entropy satisfies, $H_e(R, \rho^S) \leq H(p) \leq \log N$, which can be proved as follows

$$\begin{aligned} H_e(R, \rho^S) &= H\left(\sum_i p_i |\phi_i\rangle \langle \phi_i|^{SZ}\right) \\ &\leq H(p) \leq \log N, \end{aligned} \quad (7)$$

where in the first line, we have used $(u_i \otimes \mathbb{I}^Z)|\psi\rangle^{SZ} = |\phi_i\rangle^{SZ}$. For the example of qubit maximally coherent state, the entropy exchange is equal to one bit which is equal to $H(p = 1/2) = 1$ as obtained preceding paragraph. Similarly, for maximally coherent qudit state, the entropy exchange of the map in Eq. (5), is given by $\log_2 d$ which is different than $H(p = 1/d^2) = 2 \log_2 d$ of preceding paragraph. Also, note that in the examples that we have considered, we found maps that fully randomize the states. This may involve more cost than the required one if we want to decohere the state. This is because, the randomization operation not only erases the coherence but also the information contained in the system. Next we define general randomizing map which can decohere any

system and then ϵ -decohering map that decoheres any state with small error $\epsilon > 0$.

Randomizing map:– Let the randomization be achieved by an ensemble of incoherent unitaries $\{p_i, u_i\}_{i=1}^N$. We associate the map

$$\mathcal{R} : \rho \mapsto \sum_{i=1}^N p_i u_i^d \rho u_i^{d\dagger}, \quad (8)$$

to the ensemble of these incoherent unitaries. We will be calling this class of incoherent completely positive trace preserving (ICPTP) maps on system S as the ‘‘incoherent unitary randomizing’’ (IUR) maps. Let us also define ϵ -decohering map as follows:

ϵ -decohering IUR map:– We say that an IUR map \mathcal{R} acting on a state ρ , is ϵ -decohering, if there exists an incoherent state τ such that

$$\|\mathcal{R}(\rho) - \tau\|_1 \leq \epsilon, \quad (9)$$

where $\|\cdot\|_1$ is the trace norm [51, 52] and for a matrix A , the trace norm is defined as $\|A\|_1 = \text{Tr} \sqrt{A^\dagger A}$. With these definitions in hand, we now present our results in the next section.

III. COST OF ERASING QUANTUM COHERENCE

In this section, we prove the main result of our work. We will be concerned with asymptotic cases of the randomization procedure. But before going to the asymptotic case, let us consider the single copy case. Consider any CPTP map Υ that decoheres the system in any state ρ and maps it to some incoherent state $\rho_I = \sum_a p_a |a\rangle \langle a|$, where $\{|a\rangle\}$ is the fixed reference basis and $\{p_a\}$ defines a probability distribution, i.e.,

$$\rho \rightarrow \rho_I = \Upsilon[\rho]. \quad (10)$$

The entropy exchange of this map is given by $H_e(\Upsilon, \rho) = H\left((\Upsilon \otimes \mathbb{I}^Z)[|\psi\rangle \langle \psi|^{SZ}]\right)$, where Z is a reference system used to purify the state ρ . Now from monotonicity of mutual information, i.e., $I(\Upsilon[\rho^{SZ}]) \leq I(\rho^{SZ})$, we have

$$H_e(\Upsilon, \rho) \geq H(\rho_I) - H(\rho). \quad (11)$$

The minimum exchange entropy can be defined as $H_e^{\min} = \min_{\{p_a\}} H(\rho_I) - H(\rho)$. If this minimum is achieved for $\rho_I = \rho^d$, where ρ^d is the diagonal part of ρ in the reference basis, then $H_e^{\min} = C_r(\rho)$. In the asymptotic limit we show that this is *exactly* the case even if the CPTP map decoheres the state ρ with some nonzero small error. Now, we state our main result as the following theorem.

Theorem: The quantum coherence of a system in a state ρ , as measured by the minimal amount of noise that is to be added in order to turn it into an incoherent state, described as the coherence of erasure of ρ , is relative entropy of coherence $C_r(\rho)$,

in the asymptotic limit. Mathematically, it holds that

$$\begin{aligned} & \sup_{\epsilon > 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \min \left\{ H_\epsilon(\mathcal{R}, \rho^{\otimes n}) : \mathcal{R} \text{ } \epsilon\text{-decohering} \right\} \\ &= \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \min \left\{ \log N : \mathcal{R} \text{ } \epsilon\text{-decohering} \right\} \\ &= C_r(\rho). \end{aligned} \quad (12)$$

Proof. The proof of the theorem follows from the following two lemmas.

Lemma 1: Consider any IUR map \mathcal{R} on the n copies of the system S^n as

$$\mathcal{R} : \rho^{\otimes n} \mapsto \sum_{i=1}^N p_i U_i^L \rho^{\otimes n} U_i^{L\dagger}, \quad (13)$$

which ϵ -decoheres $\rho^{\otimes n}$. Then, the amount of entropy that is injected in the system is lower bounded as

$$H_\epsilon(\mathcal{R}, \rho^{\otimes n}) \geq n[C_r(\rho) - \epsilon \log d - H_2(\epsilon)], \quad (14)$$

where $C_r(\rho)$ is the relative entropy of coherence for the state ρ as given in Eq. (1) and $H_2(\epsilon) = -\epsilon \ln \epsilon - (1 - \epsilon) \ln(1 - \epsilon)$ is the binary Shannon entropy. In the asymptotic limit, the minimum entropy exchange, i.e., the minimum cost for erasing coherence, is given by

$$\sup_{\epsilon > 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \min \left\{ H_\epsilon(\mathcal{R}, \rho^{\otimes n}) : \mathcal{R} \text{ } \epsilon\text{-decohering} \right\} = C_r(\rho). \quad (15)$$

Proof. First of all, define

$$R_D := \mathcal{P}(\mathcal{R}[\rho^{\otimes n}]) = \sum_k \Pi_k \mathcal{R}[\rho^{\otimes n}] \Pi_k, \quad (16)$$

where $\{\Pi_k\}$ are the projectors on the product subspaces written in the reference basis for the n copies of the system. Any incoherent state under the projective measurement in the reference basis remain intact. Now utilizing the monotonicity of the trace norm under CPTP maps [51, 52], we have

$$\begin{aligned} \|R_D - \tau\|_1 &= \|\mathcal{P}(\mathcal{R}[\rho^{\otimes n}]) - \mathcal{P}(\tau)\|_1 \\ &\leq \|\mathcal{R}[\rho^{\otimes n}] - \tau\|_1 \leq \epsilon, \end{aligned} \quad (17)$$

where in the last line we have used the fact that the IUR map \mathcal{R} is an ϵ -decohering map. Now consider the following quantity

$$\|\mathcal{R}[\rho^{\otimes n}] - R_D\|_1 \leq \|\mathcal{R}[\rho^{\otimes n}] - \tau\|_1 + \|\tau - R_D\|_1 \leq 2\epsilon, \quad (18)$$

where we have used the triangle inequality for the trace distance and made use of Eq. (17) together with the fact that the IUR map \mathcal{R} is an ϵ -decohering map. Now, since $\|\mathcal{R}[\rho^{\otimes n}] - R_D\|_1 \leq 2\epsilon$, in the worst case with $\|\mathcal{R}[\rho^{\otimes n}] - R_D\|_1 = 2\epsilon$, from the Fannes-Audenaert inequality [53] (see appendix A), we have

$$\begin{aligned} |H(\mathcal{R}[\rho^{\otimes n}]) - H(R_D)| &\leq \epsilon \ln(d^n - 1) + H_2(\epsilon) \\ &\leq \epsilon n \log d + H_2(\epsilon), \end{aligned} \quad (19)$$

where in the last line we have used $\ln(d^n - 1) \leq n \log d$ and $H_2(\epsilon) = -\epsilon \ln \epsilon - (1 - \epsilon) \ln(1 - \epsilon)$. Noting the fact that R_D is the diagonal part of $\mathcal{R}[\rho^{\otimes n}]$ and $H(R_D) \geq H(\mathcal{R}[\rho^{\otimes n}])$, we have

$$H(\mathcal{R}[\rho^{\otimes n}]) \geq H(R_D) - n\epsilon \log d - H_2(\epsilon). \quad (20)$$

Here, we pause to look at entropy of R_D more closely. The incoherent unitary operations cannot change the diagonal parts of any density matrix except permuting the diagonal elements (of course they can change phases of off diagonal terms). This can be seen from the fact that any incoherent unitary can be written as a product of a unitary diagonal matrix and a permutation matrix, i.e., $U^L = V\Pi$. Therefore, we have $U^L \rho U^{L\dagger} = V \sum_{ij} \rho_{ij} |\Pi(i)\rangle \langle \Pi(j)| V^\dagger$. In the following, a superscript d on a state ρ will mean the diagonal part of the density matrix in the fixed product reference basis. Now the diagonal part of the density matrix $U^L \rho U^{L\dagger}$ is given by

$$\begin{aligned} (U^L \rho U^{L\dagger})^d &= \sum_l \langle l| V \sum_{ij} \rho_{ij} |\Pi(i)\rangle \langle \Pi(j)| V^\dagger |l\rangle \langle l| \\ &= \sum_i \rho_{\Pi(i)\Pi(i)} |\Pi(i)\rangle \langle \Pi(i)|. \end{aligned} \quad (21)$$

Therefore, we have $H((U^L \rho U^{L\dagger})^d) = H(\rho^d)$. Making use of this fact for R_D , we have

$$\begin{aligned} H(R_D) &\geq \sum_i p_i H\left(\left(U_i^L \rho^{\otimes n} U_i^{L\dagger}\right)^d\right) \\ &= \sum_i p_i H(\rho^{d\otimes n}) = nH(\rho^d). \end{aligned} \quad (22)$$

From the Eq. (20), we have

$$\begin{aligned} H(\mathcal{R}[\rho^{\otimes n}]) &\geq nH(\rho^d) - n\epsilon \log d - H_2(\epsilon) \\ &\geq n[H(\rho^d) - \epsilon \log d - H_2(\epsilon)], \end{aligned} \quad (23)$$

where in the last line, we have used $-H_2(\epsilon) \geq -nH_2(\epsilon)$. Now, we come to the question of finding the cost of randomizing operation, i.e., the entropy that we have injected in the system. For this (as in the definition), we will consider the purification of ρ which is given by ψ such that $\rho^{\otimes n} = \text{Tr}_Z(|\psi\rangle \langle \psi|^{\otimes n})$. Let us define

$$\Omega_{S^n Z^n} := (\mathbb{I}_Z^{\otimes n} \otimes \mathcal{R})[|\psi\rangle \langle \psi|^{\otimes n}]. \quad (24)$$

Since, \mathcal{R} does not act on the reference system Z , $H(\Omega_{Z^n}) = H(\text{Tr}_S(|\psi\rangle \langle \psi|^{\otimes n})) = H(\rho^{\otimes n}) = nH(\rho)$. Now,

$$\begin{aligned} H(\Omega_{S^n Z^n}) &\geq H(\Omega_{S^n}) - H(\Omega_{Z^n}) \\ &\geq H(\mathcal{R}[\rho^{\otimes n}]) - nH(\rho), \end{aligned} \quad (25)$$

where in the first line, we have made use of the Araki-Lieb inequality [51, 52, 54]. Using Eq. (23) in the above equation, we get

$$\begin{aligned} H(\Omega_{S^n Z^n}) &\geq n[H(\rho^d) - H(\rho) - \epsilon \log d - H_2(\epsilon)] \\ &= n[C_r(\rho) - \epsilon \log d - H_2(\epsilon)]. \end{aligned} \quad (26)$$

Therefore, in the asymptotic limit, the minimal entropy exchange is given as in Eq. (15). This completes the proof of the lemma 1.

Next we consider the question of cost of erasing coherence while the amount of noise injected in the system is quantified by $\log N$, where N is the number of unitaries in the ensemble comprising the ϵ -decohering map.

Lemma 2: For any state ρ and $\epsilon > 0$ there exists, for all sufficiently large n , a map

$$\mathcal{R} : \sigma \mapsto \frac{1}{N} \sum_{i=1}^N U_i \sigma U_i^\dagger \quad (27)$$

on system with U_i being unitary operators on the system, which ϵ -decoheres it, and with

$$\log N \leq n(C_r(\rho) + \epsilon), \quad (28)$$

where $C_r(\rho)$ is the relative entropy of coherence of the state ρ . In the asymptotic limit, the minimal amount of noise as quantified by $\log N$, that is injected in the system is given by

$$\sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \min \{\log N : \mathcal{R} \text{ } \epsilon\text{-decohering}\} = C_r(\rho). \quad (29)$$

Proof: Let us consider n copies of the system in the state ρ . Also, consider a typical projector Π that projects the system onto its typical subspace. Let $\tilde{\rho} = \Pi \rho^{\otimes n} \Pi$. By definition of the typical projector, we have $\text{Tr}(\Pi \rho^{\otimes n}) \geq (1 - \epsilon)$. Therefore, using the ‘‘gentle operator lemma’’ [52], we have

$$\|\rho^{\otimes n} - \tilde{\rho}\|_1 \leq 2\sqrt{\epsilon}. \quad (30)$$

Now consider an ensemble of unitaries with some probability density function $p(dU)$, i.e. $\{U, p(dU)\}$ such that, for any state γ on the typical subspace of $\rho^{\otimes n}$, it yields

$$\int_U p(dU) U \gamma U^\dagger = \frac{1}{D} \mathcal{I}_\Pi, \quad (31)$$

where $D = 2^{n(H(\rho) - \epsilon)}$ and \mathcal{I}_Π is the identity supported on the typical subspace of the system. Therefore, we have

$$\int_U p(dU) U \tilde{\rho} U^\dagger = \frac{1}{D} \mathcal{I}_\Pi := \tau. \quad (32)$$

Now, using $D_d = 2^{n(H(\rho^d) + \epsilon)}$, we have

$$\tau = \frac{1}{D} \mathcal{I}_\Pi \geq \frac{1}{D_d} \mathcal{I}_\Pi. \quad (33)$$

Then, using the ‘‘operator Chernoff bound’’ [55, 56] (see also appendix A), we show that we can select a subensemble of these unitaries which suffices the approximation. To this end, we will consider the unitaries as random operators with the distribution $p(dU)$, and define the random operators as

$$X := DU \tilde{\rho} U^\dagger. \quad (34)$$

Here $X \geq 0$. Using $\tilde{\rho} \leq \Pi/D$ (see Eq. (A10)), we have

$$X = DU \tilde{\rho} U^\dagger \leq U \Pi U^\dagger \leq \mathcal{I}. \quad (35)$$

Now, the average value $\mathbb{E}X$ of the random operator X is given by

$$\begin{aligned} \mathbb{E}X &= D \int_U p(dU) U \tilde{\rho} U^\dagger \\ &= D\tau \geq \frac{D}{D_d} \mathcal{I}_\Pi = 2^{-n(C_r(\rho) + 2\epsilon)} \mathcal{I}_\Pi, \end{aligned} \quad (36)$$

where $C_r(\rho)$ is the relative entropy of coherence of the state ρ . If X_1, \dots, X_N , where $X_i = DU_i \tilde{\rho} U_i^\dagger$ ($i = 1, \dots, N$), are N independent realizations of X , then using the operator Chernoff bound (see Eq. (A12) of appendix A), we have

$$\begin{aligned} \Pr \left((1 - \epsilon) \mathbb{E}X \leq \frac{1}{N} \sum_{i=1}^N X_i \leq (1 + \epsilon) \mathbb{E}X \right) \\ \geq 1 - 2 \dim(\Pi) \exp \left[-\frac{N\epsilon^2}{4 \ln 2} 2^{-n(C_r(\rho) + 2\epsilon)} \right]. \end{aligned} \quad (37)$$

For $N = 2^{n(C_r(\rho) + 3\epsilon)}$ or higher, we have the corresponding probability on LHS of Eq. (37) nonzero for sufficiently large n . For this case, we have

$$(1 - \epsilon) \mathbb{E}X \leq \frac{1}{N} \sum_{i=1}^N X_i \leq (1 + \epsilon) \mathbb{E}X. \quad (38)$$

This can be recast as

$$\left\| \frac{1}{N} \sum_{i=1}^N U_i \tilde{\rho} U_i^\dagger - \tau \right\|_1 \leq \epsilon. \quad (39)$$

Now, we have

$$\begin{aligned} \left\| \frac{1}{N} \sum_{i=1}^N U_i \rho^{\otimes n} U_i^\dagger - \tau \right\|_1 &\leq \epsilon + \left\| \frac{1}{N} \sum_{i=1}^N U_i (\rho^{\otimes n} - \tilde{\rho}) U_i^\dagger \right\|_1 \\ &\leq \epsilon + \|\rho^{\otimes n} - \tilde{\rho}\|_1 \\ &\leq \epsilon + 2\sqrt{\epsilon}. \end{aligned} \quad (40)$$

Therefore, there indeed exists decohering map R that $(\epsilon + 2\sqrt{\epsilon})$ -decoheres any state with, $N = 2^{n(C_r(\rho) + 3\epsilon)} \leq 2^{n(C_r(\rho) + \epsilon + 2\sqrt{\epsilon})}$, i.e., $\log N \leq n(C_r(\rho) + \epsilon + 2\sqrt{\epsilon})$. Thus, in the asymptotic limit, the minimal cost of erasing coherence is given by Eq. (29). This concludes the proof of the Lemma 2. Now, combining Lemma 1 and 2 completes the proof of the theorem. Thus, our result provides an operational interpretation of the coherence measure of the newly developed resource theory of coherence [19], i.e., the relative entropy of coherence is nothing but the minimum amount of the noise that has to be added to the system to erase the coherence. Moreover, our result is robust, that is if we allow for nonzero error in the erasing process, it still gives the same answer in the asymptotic limit.

In an independent work, Winter and Yang [57] have shown that the relative entropy of coherence, emerges as the asymptotic rate at which we can distill maximally coherent states. This is very satisfying as the same quantity, namely, the relative entropy of coherence comes up from two (apparently) completely different tasks such as the erasure and distillation of coherence. Thus, results of [57] and ours complement each other, pointing to $C_r(\rho)$ as a bonafide operational measure of quantum coherence.

IV. CONNECTION TO LANDAUER'S ERASURE

Erasure of information has a long history [31–33, 36, 37, 58–65]. In classical or quantum computer when we erase information, we need to pay a price. This is the Landauer erasure principle which says that erasure of a single bit (or qubit) of information costs one bit of entropy. In general, when we erase a string of qubits we reduce the entropy of a quantum register. That inevitably increases the entropy of its surroundings, leading to generation of heat.

Can there be a direct connection between erasure of information and erasure of quantum coherence? First, note that these two operations are in general not the same. An irreversible operation may lead to an incoherent state, thus erasing all the coherence yet there can be some information left in the final state. However, if we erase information we bleach out everything, including that of the quantum coherence. Here, we make a formal connection between the erasure of quantum coherence and erasure of information. Furthermore, we argue that the energy cost of keeping the state coherent, in the asymptotic limit, is given by $k_B T C_r(\rho)$ per copy.

The erasing of the coherence requires the erasing of the information acquired by the memory in order to implement the random incoherent unitary map. This can be seen as follows. Let the states of the memory system be described by $\{|k\rangle\}_{k=1}^N$, which occur according to probability distribution $\{p_k\}$. We implement the random incoherent unitaries based on the random bits of the memory. This step does not change the coherence. But if we choose to forget the memory bit, as is necessary to implement the random incoherent unitary channel, the memory bit acquires some (subjective) information as given by the final state of the memory system

$$\Omega^M = \sum_{k=1}^N p_k |k\rangle \langle k|^M. \quad (41)$$

Therefore, we need to erase this state of the memory (source of subjective information [36]) and this will cost work. Based on the previous studies the work cost of erasing [36, 38, 63] this state is always at least $k_B T H(\Omega^M) = k_B T H(p)$ implying that the work W_{er} required to erase the coherence, satisfies

$$W_{\text{er}} \geq k_B T H(p) \geq k_B T H_e(R, \rho^S). \quad (42)$$

Therefore, the minimum amount of work that is necessary to erase the quantum coherence of the system is equal to $k_B T C_r(\rho^S)$.

We can also give an intuitive argument as to why the relative entropy of coherence is equal to the minimal noise that needs to be injected into the system as well as the cost of maintaining the state coherent. This can be seen by combining the Schumacher compression theorem with the Landauer erasure principle. If we are given n -copies of a quantum system in state ρ , then in the asymptotic limit, the best possible compression one can achieve is the Schumacher compression [66, 67]. The information content of n -copies with arbitrarily good fidelity resides in a Hilbert space of dimension $2^{nS(\rho)}$ which is the dimension of the typical subspace of $\rho^{\otimes n}$. When

each of the system decoheres to its diagonal state, the dimension of the typical subspace increases as given by $2^{nS(\rho^d)}$. This suggests that effectively, $N_e = n(S(\rho^d) - S(\rho))$ number of extra qubits of quantum information needs to be injected to the system to decohere the state. Thus, the erasure of coherence needs $C_r(\rho)$ bits of entropy per copy. Conversely, we can ask how much does it cost to keep the state coherent? This can be achieved by throwing away N_e extra qubits from n -copies of the decohered state ρ^d which is equivalent to $nS(\rho^d)$ qubits in the typical subspace. Now using the Landauer principle, if the erasure process takes place at a temperature T , then to erase $S(\rho^d) - S(\rho)$ number of qubits effectively, we need to spend $k_B T (S(\rho^d) - S(\rho)) = k_B T C_r(\rho)$ amount of energy per copy. Physically, this suggests that $k_B T C_r(\rho)$ can also be interpreted as the amount of energy required to keep a state coherent. Further exploration along these lines will be the subject of future work.

V. CONCLUSION AND DISCUSSION

To conclude, we have provided an operational interpretation to quantum coherence as measured by the relative entropy of coherence, in terms of the amount of noise that is to be injected in a quantum system in order to fully decohere it, in the asymptotic limit. This provides thermodynamic cost of erasing quantum coherence. Moreover, our result is robust, i.e., if we allow for nonzero error in the erasing process, it still gives the same answer in the asymptotic limit. The resource theory of coherence starts with the premise that the allowed operations are the incoherent ones and the free states are the incoherent states, and proposes the relative entropy of coherence as a valid measure. However, its operational interpretation was lacking. Our results imply that with the above restrictions, the quantifier of coherence has an explicit operational significance.

Erasure of information and therefore erasure of coherence has deep connections with physics and information science. When we erase information the final state is reset to a fixed state that can be used again in quantum memory, thus bleaching out all the information. Therefore, the erasure of quantum coherence is in general different from the erasure of information. We may have a physical process where we end up with a diagonal state in some chosen basis, leading to erasure of coherence yet that may not correspond to erasure of information. However, erasure of information or hiding of information necessarily means that the system loses everything including the coherence. Here, in this work, we have argued that the erasure of coherence and cost of keeping coherence of any state, in the asymptotic limit, is given by the relative entropy of coherence. It will be very interesting to further explore the relation between no-hiding theorem and coherence erasure. This will be the subject of future work. We hope that our result may provide deep insights into the nature of coherence and interplay of information within the realm of quantum information and thermodynamics.

ACKNOWLEDGMENTS

Authors thank Andreas Winter for the useful remarks on the manuscript. U.S. and A.M. acknowledge the research fellowship of Department of Atomic Energy, Government of India.

-
- [1] C. A. Rodríguez-Rosario, T. Frauenheim, and A. Aspuru-Guzik, [arXiv:1308.1245](#).
- [2] M. Horodecki and J. Oppenheim, *Nature Communications* **4**, 2059 (2013).
- [3] P. Skrzypczyk, A. J. Short, and S. Popescu, *Nature Communications* **5**, 4185 (2014).
- [4] V. Narasimhachar and G. Gour, [arXiv:1409.7740v2](#).
- [5] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, *Proceedings of the National Academy of Sciences* **112**, 3275 (2015).
- [6] M. Lostaglio, D. Jennings, and T. Rudolph, *Nature Communications* **6** (2015).
- [7] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, *Phys. Rev. X* **5**, 021001 (2015).
- [8] B. Gardas and S. Deffner, [arXiv:1503.03455](#).
- [9] A. Misra, U. Singh, M. N. Bera, and A. K. Rajagopal, [arXiv:1505.06980](#).
- [10] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, [arXiv:1505.07835](#).
- [11] M. B. Plenio and S. F. Huelga, *New J. Phys.* **10**, 113019 (2008).
- [12] P. Reberntrost, M. Mohseni, and A. Aspuru-Guzik, *J. Phys. Chem. B* **113**, 9942 (2009).
- [13] S. Lloyd, *J. Phys.: Conf. Ser.* **302**, 012037 (2011).
- [14] C.-M. Li, N. Lambert, Y.-N. Chen, G.-Y. Chen, and F. Nori, *Sci. Rep.* **2**, 885 (2012).
- [15] S. Huelga and M. Plenio, *Contemporary Physics* **54**, 181 (2013).
- [16] F. Levi and F. Mintert, *New J. Phys.* **16**, 033007 (2014).
- [17] G. Gour and R. W. Spekkens, *New Journal of Physics* **10**, 033023 (2008).
- [18] I. Marvian and R. W. Spekkens, *Phys. Rev. A* **90**, 062110 (2014).
- [19] T. Baumgratz, M. Cramer, and M. B. Plenio, *Phys. Rev. Lett.* **113**, 140401 (2014).
- [20] T. R. Bromley, M. Cianciaruso, and G. Adesso, *Phys. Rev. Lett.* **114**, 210401 (2015).
- [21] D. Girolami, *Phys. Rev. Lett.* **113**, 170401 (2014).
- [22] I. Marvian and R. W. Spekkens, *Nature Communications* **5**, 3821 (2014).
- [23] Z. Xi, Y. Li, and H. Fan, [arXiv:1408.3194](#).
- [24] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, [arXiv:1502.05876](#).
- [25] L.-H. Shao, Z. Xi, H. Fan, and Y. Li, *Phys. Rev. A* **91**, 042120 (2015).
- [26] D. P. Pires, L. C. Céleri, and D. O. Soares-Pinto, *Phys. Rev. A* **91**, 042330 (2015).
- [27] S. Du, Z. Bai, and Y. Guo, *Phys. Rev. A* **91**, 052120 (2015).
- [28] Y. Yao, X. Xiao, L. Ge, and C. P. Sun, [arXiv:1506.01773](#).
- [29] M. N. Bera, T. Qureshi, M. A. Siddiqui, and A. K. Pati, [arXiv:1503.02990](#).
- [30] U. Singh, M. N. Bera, H. S. Dhar, and A. K. Pati, *Phys. Rev. A* **91**, 052115 (2015).
- [31] R. Landauer, *IBM Journal of Research and Development* **5**, 183 (1961).
- [32] C. H. Bennett, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **34**, 501 (2003).
- [33] M. B. Plenio and V. Vitelli, *Contemporary Physics* **42**, 25 (2001).
- [34] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **89**, 180402 (2002).
- [35] M. Horodecki, P. Horodecki, and J. Oppenheim, *Phys. Rev. A* **67**, 062104 (2003).
- [36] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, *Open Systems & Information Dynamics* **11**, 205 (2004).
- [37] L. d. Rio, J. Aberg, R. Renner, O. Dahlsten, and V. Vedral, *Nature* **474**, 61 (2011).
- [38] B. Groisman, S. Popescu, and A. Winter, *Phys. Rev. A* **72**, 032317 (2005).
- [39] M. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen, and B. Synak-Radtke, *Phys. Rev. A* **71**, 062307 (2005).
- [40] A. K. Pati, [arXiv:1208.4804](#).
- [41] F. Buscemi, *New Journal of Physics* **11**, 123002 (2009).
- [42] E. Wakakuwa, A. Soeda, and M. Murao, [arXiv:1504.05805](#).
- [43] M. Mosca, A. Tapp, and R. de Wolf, [arXiv:quant-ph/0003101](#).
- [44] P. O. Boykin and V. Roychowdhury, *Phys. Rev. A* **67**, 042317 (2003).
- [45] P. Hayden, D. Leung, P. W. Shor, and A. Winter, *Communications in Mathematical Physics* **250**, 371 (2004).
- [46] B. Schumacher, *Phys. Rev. A* **54**, 2614 (1996).
- [47] B. Schumacher and M. A. Nielsen, *Phys. Rev. A* **54**, 2629 (1996).
- [48] S. L. Braunstein and A. K. Pati, *Phys. Rev. Lett.* **98**, 080502 (2007).
- [49] J. R. Samal, A. K. Pati, and A. Kumar, *Phys. Rev. Lett.* **106**, 080401 (2011).
- [50] Y. Xiang and S.-J. Xiong, *Phys. Rev. A* **76**, 014306 (2007).
- [51] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).
- [52] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
- [53] K. M. R. Audenaert, *Journal of Physics A: Mathematical and Theoretical* **40**, 8127 (2007).
- [54] H. Araki and E. Lieb, *Communications in Mathematical Physics* **18**, 160 (1970).
- [55] R. Ahlswede and A. Winter, *IEEE Transactions on Information Theory* **48**, 569 (2002).
- [56] R. Ahlswede and A. Winter, *IEEE Transactions on Information Theory* **49**, 346 (2003).
- [57] A. Winter and D. Yang, [arXiv:1506.TODAY](#).
- [58] B. Piechocinska, *Phys. Rev. A* **61**, 062314 (2000).
- [59] A. E. Allahverdyan and T. M. Nieuwenhuizen, *Phys. Rev. E* **64**, 056117 (2001).
- [60] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **102**, 250602 (2009).
- [61] R. Dillenschneider and E. Lutz, *Phys. Rev. Lett.* **102**, 210601 (2009).
- [62] S. Hilt, S. Shabbir, J. Anders, and E. Lutz, *Phys. Rev. E* **83**,

030102 (2011).

- [63] A. Hosoya, K. Maruyama, and Y. Shikano, *Phys. Rev. E* **84**, 061117 (2011).
 [64] M. Das, *Phys. Rev. E* **89**, 032130 (2014).
 [65] J. Goold, M. Paternostro, and K. Modi, *Phys. Rev. Lett.* **114**, 060602 (2015).
 [66] R. Jozsa and B. Schumacher, *Journal of Modern Optics* **41**, 2343 (1994), <http://dx.doi.org/10.1080/09500349414552191>.
 [67] B. Schumacher, *Phys. Rev. A* **51**, 2738 (1995).
 [68] A. Winter, *IEEE Transactions on Information Theory* **45**, 2481 (1999).
 [69] T. Ogawa and H. Nagaoka, *IEEE Transactions on Information Theory* **53**, 2261 (2007).
 [70] M. Fannes, *Communications in Mathematical Physics* **31**, 291 (1973).

Appendix A: Trace distance and gentle operator lemma

For the sake of completeness, here we provide the already well known results in the context of trace distance, such as gentle operator lemma [68, 69] and Fannes-Audenaert Inequality [53]. We also state operator Chernoff bound [55, 56] and list properties of typical subspaces. All these results are quoted from [52] along with original references.

In the context of continuity of von Neuman entropy, Audenaert proved a tighter inequality than Fannes' inequality [70], which is now known as Fannes-Audenaert inequality [53] and can be stated as

Fannes-Audenaert Inequality ([53]):– For any ρ and σ with $T \equiv \frac{1}{2}\|\rho - \sigma\|_1$, the following inequality holds:

$$|H(\rho) - H(\sigma)| \leq T \log(d-1) + H_2(T), \quad (\text{A1})$$

where d is the dimension of the Hilbert space of the system in the state ρ and $H_2(T) = -T \ln T - (1-T) \ln(1-T)$ is the binary Shannon entropy. Next we state, the gentle operator lemma, which was first given in Ref. [68] and later improved in Ref. [69] (see also [52]).

Gentle operator lemma ([68, 69]):– Suppose that a measurement operator Λ ($0 \leq \Lambda \leq I$) has a high probability of detecting a subnormalised state ρ , i.e., $\text{Tr}\{\Lambda\rho\} \geq \text{Tr}\{\rho\} - \epsilon$, where $1 \geq \epsilon > 0$ and ϵ is close to zero. Then $\sqrt{\Lambda}\rho\sqrt{\Lambda}$ is $2\sqrt{\epsilon}$ -close to the original state ρ in trace distance:

$$\left\| \rho - \sqrt{\Lambda}\rho\sqrt{\Lambda} \right\|_1 \leq 2\sqrt{\epsilon}, \quad (\text{A2})$$

where $\|\sigma\|_1 = \text{Tr}\sqrt{\sigma^\dagger\sigma}$.

Typical Sequence and Typical Set:– A sequence x^n is δ -typical if its sample entropy $\overline{H}(x^n)$ is δ -close to the entropy $H(X)$ of random variable X , where this random variable is the source of the sequence. The set of all δ -typical sequences x^n is defined as the typical set $T_\delta^{X^n}$, i.e.,

$$T_\delta^{X^n} \equiv \{x^n : |\overline{H}(x^n) - H(X)| \leq \delta\}. \quad (\text{A3})$$

Now, consider a quantum state with spectral decomposition

as

$$\rho^X = \sum_x p_X(x) |x\rangle\langle x|^X. \quad (\text{A4})$$

Considering n copies of the state ρ^X , we have

$$(\rho^X)^{\otimes n} := \rho^{X^n} = \sum_{x^n} p_{X^n}(x^n) |x^n\rangle\langle x^n|^{X^n}, \quad (\text{A5})$$

where $X^n = (X_1 \dots X_n)$, $x^n = (x_1 \dots x_n)$, $p_{X^n}(x^n) = p_X(x_1) \dots p_X(x_n)$ and $|x^n\rangle = |x_1\rangle^{X_1} \otimes \dots \otimes |x_n\rangle^{X_n}$.

Typical Subspace:– The δ -typical subspace $T_{\rho,\delta}^{X^n}$ is a subspace of the full Hilbert space X_1, \dots, X_n and is spanned by states $|x^n\rangle^{X^n}$ whose corresponding classical sequences x^n are δ -typical:

$$T_{\rho,\delta}^{X^n} \equiv \text{span} \left\{ |x^n\rangle^{X^n} : x^n \in T_\delta^{X^n} \right\}. \quad (\text{A6})$$

Also, one can define a typical projector, which is projector onto the typical subspace, as

$$\Pi_{\rho,\delta}^{X^n} \equiv \sum_{x^n \in T_\delta^{X^n}} |x^n\rangle\langle x^n|^{X^n}. \quad (\text{A7})$$

Properties of typical subspaces:– (a) The probability that the quantum state ρ^{X^n} is in the typical subspace $T_{\rho,\delta}^{X^n}$ approaches one as n becomes large:

$$\forall \epsilon > 0 \quad \text{Tr} \left\{ \Pi_{\rho,\delta}^{X^n} \rho^{X^n} \right\} \geq 1 - \epsilon \quad \text{for sufficiently large } n, \quad (\text{A8})$$

where $\Pi_{\rho,\delta}^{X^n}$ is the typical subspace projector. (b) The dimension $\dim(T_{\rho,\delta}^{X^n})$ of the δ -typical subspace satisfies

$$\forall \epsilon > 0 \quad (1 - \epsilon) 2^{n(H(X) - \delta)} \leq \text{Tr} \left\{ \Pi_{\rho,\delta}^{X^n} \right\} \leq 2^{n(H(X) + \delta)}, \quad (\text{A9})$$

for sufficiently large n . (c) For all n the operator $\Pi_{\rho,\delta}^{X^n} \rho^{X^n} \Pi_{\rho,\delta}^{X^n}$ satisfies

$$2^{-n(H(X) + \delta)} \Pi_{\rho,\delta}^{X^n} \leq \Pi_{\rho,\delta}^{X^n} \rho^{X^n} \Pi_{\rho,\delta}^{X^n} \leq 2^{-n(H(X) - \delta)} \Pi_{\rho,\delta}^{X^n}. \quad (\text{A10})$$

Operator Chernoff Bound ([55, 56], see also [52]):– Let X_1, \dots, X_n ($\forall m \in [n] : 0 \leq X_m \leq I$) be n independent and identically distributed random variables with values in the algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators on some Hilbert space \mathcal{H} . Let \overline{X} denote the sample average of the n random variables: $\overline{X} = \frac{1}{n} \sum_{m=1}^n X_m$. Suppose that for each operator X_m

$$\mathbb{E}_X \{X_m\} \geq aI, \quad (\text{A11})$$

where $a \in (0, 1)$. Then for every ϵ where $0 < \epsilon < 1/2$ and $(1 + \epsilon)a \leq 1$, the probability that the sample average \overline{X} lies inside the operator interval $[(1 \pm \epsilon)\mathbb{E}_X \{X_m\}]$ is bounded as

$$\begin{aligned} \Pr \left\{ (1 - \epsilon) \mathbb{E}_X \{X_m\} \leq \overline{X} \leq (1 + \epsilon) \mathbb{E}_X \{X_m\} \right\} \\ \geq 1 - 2 \dim(\mathcal{H}) \exp \left(-\frac{n\epsilon^2 a}{4 \ln 2} \right). \end{aligned} \quad (\text{A12})$$