

# Purification Theorem

## Formal statement

Let  $\mathcal{H}(A)$  be a finite-dimensional complex Hilbert space, and let  $\rho_A$  be a density operator on  $\mathcal{H}(A)$ . Thus

$$\rho_A \geq 0, \quad \text{Tr } \rho_A = 1.$$

The purification theorem says that there exists another finite-dimensional Hilbert space  $\mathcal{H}(R)$ , called a reference system, and a normalized pure state

$$|\psi\rangle_{AR} \in \mathcal{H}_A \otimes \mathcal{H}_R$$

such that

$$\rho_A = \text{Tr}_R (|\psi\rangle\langle\psi|_{AR}).$$

In words, every mixed state can be regarded as the reduced state, or marginal state, of a larger pure state.

Moreover, if

$$\text{rank}(\rho_A) = r,$$

then a reference system of dimension  $r$  is sufficient, and no smaller reference system can purify  $\rho_A$ . Thus the minimal purification dimension is exactly the rank of the mixed state.

This theorem is one of the basic bridges between density operators and entanglement. It is usually taught together with the Schmidt decomposition, because the proof is almost the Schmidt decomposition read backwards. Standard presentations can be found in Nielsen and Chuang, Preskill, Watrous, and Wilde.

## Proof

The proof begins with the spectral theorem. Since  $\rho_A$  is Hermitian and positive semidefinite, it has an orthonormal eigenbasis on its support. Therefore we can write

$$\rho_A = \sum_{i=1}^r \lambda_i |i\rangle\langle i|_A,$$

where

$$\lambda_i > 0, \quad \sum_{i=1}^r \lambda_i = 1,$$

and  $r = \text{rank}(\rho_A)$ . The zero eigenvalues do not need to be included.

Now introduce a reference system R with Hilbert space  $\mathcal{H}(R)$  of dimension at least r. Choose an orthonormal set

$$\{|i\rangle_R : i = 1, \dots, r\}$$

inside  $\mathcal{H}(R)$ . Define

$$|\psi\rangle_{AR} = \sum_{i=1}^r \sqrt{\lambda_i} |i\rangle_A |i\rangle_R.$$

This state is normalized because

$$\langle\psi|\psi\rangle = \sum_{i,j} \sqrt{\lambda_i \lambda_j} \langle j|i\rangle_A \langle j|i\rangle_R = \sum_i \lambda_i = 1.$$

Now compute the density operator of the joint pure state:

$$|\psi\rangle\langle\psi|_{AR} = \sum_{i,j} \sqrt{\lambda_i \lambda_j} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_R.$$

Taking the partial trace over R, we use

$$\text{Tr}_R(|i\rangle\langle j|_R) = \langle j|i\rangle_R = \delta_{ij}.$$

Therefore

$$\begin{aligned} \text{Tr}_R(|\psi\rangle\langle\psi|_{AR}) &= \sum_{i,j} \sqrt{\lambda_i\lambda_j} |i\rangle\langle j|_A \text{Tr}(|i\rangle\langle j|_R) \\ &= \sum_{i,j} \sqrt{\lambda_i\lambda_j} |i\rangle\langle j|_A \delta_{ij} \\ &= \sum_i \lambda_i |i\rangle\langle i|_A \\ &= \rho_A. \end{aligned}$$

Thus  $|\psi\rangle_{AR}$  is a purification of  $\rho_A$ .

It remains to explain the minimal dimension statement. Suppose  $|\varphi\rangle_{AR}$  is any purification of  $\rho_A$ . Since  $|\varphi\rangle_{AR}$  is a pure bipartite state, it has a Schmidt decomposition

$$|\varphi\rangle_{AR} = \sum_{k=1}^s \sqrt{\mu_k} |u_k\rangle_A |v_k\rangle_R,$$

where  $s \leq \dim \mathcal{H}(R)$ . Tracing out  $R$ , we get

$$\rho_A = \sum_{k=1}^s \mu_k |u_k\rangle\langle u_k|_A.$$

Hence  $\rho_A$  has rank  $s$ , the Schmidt rank of  $|\varphi\rangle_{AR}$ . Therefore

$$\text{rank}(\rho_A) = s \leq \dim \mathcal{H}_R.$$

So any purifying system must have dimension at least  $\text{rank}(\rho_A)$ . Since the construction above works with  $\dim \mathcal{H}(R) = r$ , the minimal reference dimension is exactly  $r$ . This completes the proof.

## Operational meaning

The purification theorem says that mixedness can always be represented as missing information about a larger pure state.

This sentence must be read carefully. The theorem does not say that every mixed state really came from a larger pure state in a laboratory. Rather, it says that the mathematics of every mixed state is compatible with such a representation. Whenever we write

$$\rho_A = \text{Tr}_R(|\psi\rangle\langle\psi|_{AR}),$$

we are saying that system A can be understood as part of a larger system AR, where the total state is pure but the reference system R is inaccessible, ignored, or deliberately abstract.

The reference system R is sometimes a real physical environment. For example, a photon may be entangled with an unobserved mode, or an atom may be correlated with radiation that escaped into the environment. But in quantum information theory, R is often a mathematical reference system. It is a bookkeeping device that remembers how much information is missing from A.

The most important mental image is this: a mixed state on A can be purified by giving it a partner R that carries the complementary information. If A is mixed, then in one possible purification A is entangled with R. The eigenvalues of  $\rho(A)$  become the Schmidt probabilities of the pure state  $|\psi\rangle_{AR}$ . Thus the purification theorem converts local mixedness into bipartite entanglement.

This is why purification is so useful. Many arguments about mixed states are hard because mixed states are convex combinations of pure states in many possible ways. Purification replaces the mixed state by a pure state on a larger system, where we can use pure-state tools such as inner products, Schmidt decomposition, unitary dynamics, and entanglement entropy.

## How to use the theorem

The practical recipe is simple. To purify  $\rho(A)$ , first diagonalize it:

$$\rho_A = \sum_i \lambda_i |i\rangle\langle i|_A.$$

Then create a reference system R with matching orthonormal labels  $|i\rangle_R$ , and write

$$|\psi\rangle_{AR} = \sum_i \sqrt{\lambda_i} |i\rangle_A |i\rangle_R.$$

This is called the canonical purification relative to the eigenbasis of  $\rho(A)$ . The word “canonical” should not be taken too strongly, because if  $\rho(A)$  has degenerate eigenvalues, the eigenbasis is not unique. Still, this construction is the standard one.

Once a purification is available, many calculations become easier. The von Neumann entropy of  $\rho(A)$  becomes the entanglement entropy of  $|\psi\rangle_{AR}$ . A noisy process on  $A$  can be studied by imagining how it acts on one half of a purified state. Channel distinguishability, coherent information, decoupling, quantum error correction, and entanglement theory all use this purified viewpoint.

### Example 1: purifying a diagonal qubit state

Consider the qubit mixed state

$$\rho_A = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|, \quad 0 \leq p \leq 1.$$

A purification is

$$|\psi\rangle_{AR} = \sqrt{p}|0\rangle_A |0\rangle_R + \sqrt{1-p}|1\rangle_A |1\rangle_R.$$

Let us verify this explicitly. The joint density operator is

$$\begin{aligned} |\psi\rangle\langle\psi|_{AR} &= p|0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_R \\ &\quad + \sqrt{p(1-p)}|0\rangle\langle 1|_A \otimes |0\rangle\langle 1|_R \\ &\quad + \sqrt{p(1-p)}|1\rangle\langle 0|_A \otimes |1\rangle\langle 0|_R \\ &\quad + (1-p)|1\rangle\langle 1|_A \otimes |1\rangle\langle 1|_R. \end{aligned}$$

When we trace out  $R$ , the cross terms vanish because

$$\text{Tr}(|0\rangle\langle 1|) = 0, \quad \text{Tr}(|1\rangle\langle 0|) = 0.$$

So

$$\text{Tr}_R(|\psi\rangle\langle\psi|_{AR}) = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| = \rho_A.$$

Operationally, this purification says that the local uncertainty in A can be represented as correlation with R. If someone measured R in the basis  $\{|0\rangle, |1\rangle\}$ , they would learn which corresponding basis state appears on A. But if R is inaccessible, system A alone is described by the mixed state  $\rho(A)$ .

## Example 2: the maximally mixed qubit

The maximally mixed qubit is

$$\rho_A = \frac{I}{2}.$$

A purification is the Bell state

$$|\Phi^+\rangle_{AR} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

Tracing out R, we get

$$\text{Tr}_R(|\Phi^+\rangle\langle\Phi^+|) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{I}{2}.$$

This example is operationally important because it shows that maximal local ignorance can arise from maximal global knowledge. The state  $|\Phi^+\rangle_{AR}$  is pure; if we know the entire system AR, there is no classical uncertainty about the global state. Yet system A alone is completely mixed. The uncertainty in A is not ordinary ignorance about which pure state was secretly prepared. It can be understood as entanglement with the reference system.

This is one of the central ideas of quantum information: a subsystem of a pure entangled state can be mixed.

## Example 3: purifying a non-diagonal mixed state

Consider

$$\rho_A = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

This state is not diagonal in the computational basis. To purify it cleanly, we first diagonalize it. Its eigenvectors are

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}},$$

with eigenvalues

$$\lambda_+ = \frac{3}{4}, \quad \lambda_- = \frac{1}{4}.$$

Therefore

$$\rho_A = \frac{3}{4}|+\rangle\langle+| + \frac{1}{4}|-\rangle\langle-|.$$

A purification is

$$|\psi\rangle_{AR} = \sqrt{\frac{3}{4}}|+\rangle_A|0\rangle_R + \sqrt{\frac{1}{4}}|-\rangle_A|1\rangle_R.$$

The labels on R do not need to be  $|+\rangle$  and  $|-\rangle$ ; they only need to be orthonormal. What matters is that the reference system has one orthonormal label for each nonzero eigenvalue of  $\rho_A$ .

This example shows the correct order of operations. One should not purify a non-diagonal density matrix merely by taking square roots of its matrix entries. The right construction uses the eigenvalues and eigenvectors of  $\rho_A$ , or equivalently a square-root factorization of  $\rho_A$ .

### Example 4: different purifications of the same state

The purification of a mixed state is not unique. For the maximally mixed qubit,

$$\rho_A = \frac{I}{2},$$

the Bell state

$$|\Phi^+\rangle_{AR} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

is one purification. But so is

$$|\Phi^-\rangle_{AR} = \frac{|00\rangle - |11\rangle}{\sqrt{2}}.$$

Indeed, the relative phase disappears when we trace out R. More generally, if  $|\psi\rangle_{AR}$  purifies  $\rho(A)$ , then

$$(I_A \otimes U_R)|\psi\rangle_{AR}$$

also purifies  $\rho(A)$  for every unitary  $U(R)$  acting only on the reference system. This is because operations on a system that is later traced out cannot change the reduced state of A.

There is also a converse. If two pure states purify the same  $\rho(A)$ , then they are related by an isometry on the purifying system. If the purifying systems have the same dimension, the isometry can be taken to be a unitary. This uniqueness-up-to-isometry idea is the structural reason purification is powerful: the reference system contains freedom, but all of that freedom is outside A. Uhlmann's theorem and the Hughston-Jozsa-Wootters theorem are deeper developments of this principle.

### Example 5: ensemble decompositions and measuring the reference

A mixed state can have many ensemble decompositions. For example,

$$\frac{I}{2} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|,$$

but also

$$\frac{I}{2} = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|.$$

The purification theorem helps explain why this is possible. Start from

$$|\Phi^+\rangle_{AR} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

If the reference system R is measured in the computational basis, then system A is prepared as  $|0\rangle$  or  $|1\rangle$ , each with probability  $1/2$ . But if R is measured in the X-basis, then system A is prepared as  $|+\rangle$  or  $|-\rangle$ , again each with probability  $1/2$ , up to a harmless known correction depending on the precise Bell state convention.

Thus different ensemble decompositions of the same mixed state can be understood as different measurements on the purifying reference. This is the operational content behind the Hughston-Jozsa-Wootters theorem: all ensemble realizations of a density operator can be generated by suitable measurements on a purifying system.

This point is subtle but essential. The density matrix  $I/2$  does not remember whether it was prepared as a random mixture of  $|0\rangle, |1\rangle$  or as a random mixture of  $|+\rangle, |-\rangle$ . Those are different preparation stories for the same operational state on A. No measurement performed only on A can distinguish the stories.

## Purification and entanglement entropy

If  $|\psi\rangle_{AR}$  purifies  $\rho_A$ , then the Schmidt coefficients of  $|\psi\rangle_{AR}$  are the square roots of the nonzero eigenvalues of  $\rho_A$ . Therefore the entropy of  $\rho_A$  is the entanglement entropy between A and R:

$$S(\rho_A) = -\text{Tr}(\rho_A \log \rho_A) = -\sum_i \lambda_i \log \lambda_i.$$

If  $\rho_A$  is pure, then its rank is one, and the purification can be a product state:

$$|\psi\rangle_{AR} = |\phi\rangle_A |0\rangle_R.$$

There is no entanglement needed. If  $\rho_A$  is mixed, then every purification must be entangled across A:R. The more mixed  $\rho_A$  is, the more entanglement is required in the purification, as quantified by the entropy of the eigenvalue distribution.

This gives a useful operational interpretation of entropy. The entropy of  $\rho_A$  measures how much entanglement with a reference system is needed to regard  $\rho_A$  as part of a pure state.

## Common mistakes

A common mistake is to think that a purification is unique. It is not. The reduced state  $\rho_A$  fixes the Schmidt coefficients of any purification, but it does not fix the basis of the reference system. Acting with a unitary on  $R$  changes the purification without changing  $\rho_A$ .

Another common mistake is to interpret purification as a claim about what physically happened. If  $\rho_A$  is mixed, it may have arisen from classical random preparation, from entanglement with an environment, from noise, or from deliberate coarse-graining. The purification theorem says that all these cases can be mathematically represented by a pure state on a larger space. It does not say that the reference system is always physically present in the experimental setup.

A third mistake is to forget the minimal dimension condition. If  $\rho_A$  has rank  $r$ , then a reference system of dimension  $r$  is sufficient and necessary. For example, a full-rank qubit state needs a two-dimensional reference for a minimal purification. A pure state has rank one, so it needs no nontrivial reference system.

## Final mental image

The purification theorem says that every density matrix can be lifted into a pure state by adding a reference system. The mixed state is what remains when the reference is ignored:

$$\rho_A = \text{Tr}_R(|\psi\rangle\langle\psi|_{AR}).$$

The reference system is a mathematical holder of the missing information. When  $\rho_A$  is diagonalized as

$$\rho_A = \sum_i \lambda_i |i\rangle\langle i|,$$

the purification

$$|\psi\rangle_{AR} = \sum_i \sqrt{\lambda_i} |i\rangle_A |i\rangle_R$$

turns the eigenvalue distribution of  $\rho(A)$  into the Schmidt spectrum of a larger pure state. Thus mixedness on A becomes entanglement between A and R.

This is why purification is used everywhere in quantum information theory. It lets us replace a mixed-state problem by a pure-state problem on a larger Hilbert space. Once that replacement is made, the tools of pure-state quantum mechanics become available again.

## References

Nielsen, Michael A., and Isaac L. Chuang. Quantum Computation and Quantum Information. Cambridge University Press, 10th anniversary edition, 2010. See Section 2.5 on the Schmidt decomposition and purifications.

Preskill, John. Lecture Notes for Physics 229: Quantum Information and Computation, especially the chapters on quantum entanglement and quantum information theory.

Watrous, John. The Theory of Quantum Information. Cambridge University Press, 2018. See the early chapters on density operators, purifications, and Schmidt decompositions.

Wilde, Mark M. Quantum Information Theory. Cambridge University Press, 2nd edition, 2017. See the discussion of purified systems, entropy, and information-processing protocols.

Uhlmann, Armin. "The 'transition probability' in the state space of a -algebra." Reports on Mathematical Physics 9, no. 2 (1976): 273-279.

Hughston, Lane P., Richard Jozsa, and William K. Wootters. "A complete classification of quantum ensembles having a given density matrix." Physics Letters A 183, no. 1 (1993): 14-18.

# Document information

## Purification Theorem

---

|                      |   |
|----------------------|---|
| <b>Project</b>       | [QIT 001] Mathematical Representation Theorems  |
| <b>Document</b>      | Primary document  |
| <b>Author</b>        | mujirin   |
| <b>Verifier</b>      | Not verified  |
| <b>Downloaded</b>    | July 04, 2026 18:33 KST   |
| <b>Status</b>        | Working   |
| <b>Document link</b> | <a href="https://theorytrace.com/projects/mathematical-representation-theorems/documents/un-titled-document-a7ebff/">https://theorytrace.com/projects/mathematical-representation-theorems/documents/un-titled-document-a7ebff/</a> |