

# Schmidt Decomposition

## Formal statement

Let  $\mathcal{H}_A$  and  $\mathcal{H}_B$  be finite-dimensional complex Hilbert spaces, and let

$$|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$$

be a normalized pure bipartite state. Then there exist an integer

$$r \leq \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\},$$

positive numbers

$$\lambda_1, \dots, \lambda_r > 0, \quad \sum_{i=1}^r \lambda_i = 1,$$

and orthonormal sets

$$\{|u_1\rangle_A, \dots, |u_r\rangle_A\} \subset \mathcal{H}_A, \quad \{|v_1\rangle_B, \dots, |v_r\rangle_B\} \subset \mathcal{H}_B$$

such that

$$|\psi\rangle_{AB} = \sum_{i=1}^r \sqrt{\lambda_i} |u_i\rangle_A |v_i\rangle_B.$$

This expression is called the Schmidt decomposition of  $|\psi\rangle_{AB}$ . The positive numbers  $\sqrt{\lambda_i}$  are the Schmidt coefficients, the numbers  $\lambda_i$  are the Schmidt probabilities or Schmidt spectrum, and  $r$  is the Schmidt rank.

The notation often appears in the shorter form

$$|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i} |i\rangle_A |i\rangle_B.$$

This is convenient, but it should not be misunderstood. The two vectors both carry the same label  $i$ , but they live in different Hilbert spaces. More explicitly, one should read the expression as

$$|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i} |u_i\rangle_A |v_i\rangle_B.$$

The theorem says that every pure bipartite state can be written as a perfectly correlated superposition of orthonormal local states.

### Proof by singular value decomposition

Choose an orthonormal basis  $\{|a\rangle_A\}$  of mathcal  $H(A)$  and an orthonormal basis  $\{|\mu\rangle_B\}$  of mathcal  $H(B)$ . In these product bases, the state has the expansion

$$|\psi\rangle_{AB} = \sum_{a,\mu} M_{a\mu} |a\rangle_A |\mu\rangle_B.$$

The coefficients  $M_{(a\mu)}$  form a complex matrix  $M$ . The normalization of  $|\psi\rangle$  says

$$1 = \langle\psi|\psi\rangle = \sum_{a,\mu} |M_{a\mu}|^2.$$

Thus the squared Frobenius norm of  $M$  is one.

Now apply the ordinary singular value decomposition to the matrix  $M$ . There exist unitary matrices  $U$  and  $V$ , and nonnegative singular values  $s_1, \dots, s_r > 0$ , such that

$$M = UDV^\dagger,$$

where  $D$  is diagonal on its nonzero support and has diagonal entries  $s_i$ . Equivalently,

$$M_{a\mu} = \sum_{i=1}^r U_{ai} s_i \overline{V_{\mu i}}.$$

Define local vectors

$$|u_i\rangle_A = \sum_a U_{ai} |a\rangle_A,$$

and

$$|v_i\rangle_B = \sum_\mu \overline{V_{\mu i}} |\mu\rangle_B.$$

Because  $U$  and  $V$  are unitary, the vectors  $\{|u_i\rangle_A\}$  are orthonormal in mathcal  $H(A)$ , and the vectors  $\{|v_i\rangle_B\}$  are orthonormal in mathcal  $H(B)$ . Substituting the singular value expansion of  $M$  into the state gives

$$\begin{aligned} |\psi\rangle_{AB} &= \sum_{a,\mu} M_{a\mu} |a\rangle_A |\mu\rangle_B \\ &= \sum_{a,\mu} \sum_{i=1}^r U_{ai} s_i \overline{V_{\mu i}} |a\rangle_A |\mu\rangle_B \\ &= \sum_{i=1}^r s_i \left( \sum_a U_{ai} |a\rangle_A \right) \left( \sum_\mu \overline{V_{\mu i}} |\mu\rangle_B \right) \\ &= \sum_{i=1}^r s_i |u_i\rangle_A |v_i\rangle_B. \end{aligned}$$

Finally, since  $|\psi\rangle$  is normalized,

$$1 = \sum_{a,\mu} |M_{a\mu}|^2 = \sum_{i=1}^r s_i^2.$$

Set

$$\lambda_i = s_i^2.$$

Then  $\lambda_i > 0$ ,  $\sum_i \lambda_i = 1$ , and

$$|\psi\rangle_{AB} = \sum_{i=1}^r \sqrt{\lambda_i} |u_i\rangle_A |v_i\rangle_B.$$

This proves the theorem. The proof also tells us how to find the decomposition: write the bipartite state as a coefficient matrix and take its singular value decomposition. Watrous presents the Schmidt decomposition in exactly this vectorization-and-SVD way: a vector in  $\mathcal{X} \otimes \mathcal{Y}$  is associated with an operator, and the singular value decomposition of that operator becomes the Schmidt decomposition of the vector. Preskill likewise derives the Schmidt form by applying unitary transformations on the left and right of the coefficient matrix, identifying the Schmidt coefficients with the singular values.

## The same theorem through reduced density operators

There is another proof that is often more useful physically. From the pure state

$$|\psi\rangle_{AB} = \sum_{i=1}^r \sqrt{\lambda_i} |u_i\rangle_A |v_i\rangle_B,$$

form the density operator

$$|\psi\rangle\langle\psi|_{AB}.$$

If we trace out system B, we get

$$\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|) = \sum_{i=1}^r \lambda_i |u_i\rangle\langle u_i|_A.$$

If we trace out system A, we get

$$\rho_B = \text{Tr}_A(|\psi\rangle\langle\psi|) = \sum_{i=1}^r \lambda_i |v_i\rangle\langle v_i|_B.$$

Therefore  $\rho_A$  and  $\rho_B$  have the same nonzero eigenvalues. The zero eigenvalues may differ if  $\dim \mathcal{H}(A)$  and  $\dim \mathcal{H}(B)$  have different dimensions, but the nonzero part of the spectrum is shared.

This observation is one of the most important operational consequences of the Schmidt decomposition. A bipartite pure state may live in a large tensor-product space, but the uncertainty seen locally by A and the uncertainty seen locally by B are governed by the same probability distribution  $\{\lambda_i\}$ . Preskill explicitly emphasizes this consequence after the Schmidt decomposition: the reduced states  $\rho(A)$  and  $\rho(B)$  have the same nonzero eigenvalues.

## Operational meaning

The Schmidt decomposition says that pure bipartite entanglement has a canonical shape. If two systems A and B share a pure state, then there are local coordinate systems in which the state looks like a correlated sum

$$\sqrt{\lambda_1}|u_1\rangle_A|v_1\rangle_B + \sqrt{\lambda_2}|u_2\rangle_A|v_2\rangle_B + \dots$$

The same index  $i$  appears on both sides. This means that, in the Schmidt bases, system A and system B are perfectly correlated. If a measurement of A in the Schmidt basis gives the result  $i$ , then a measurement of B in its corresponding Schmidt basis also gives  $i$ , with probability  $\lambda_i$ .

This is the operational mental image: a pure entangled state is a coherent superposition of matched local alternatives. The probabilities  $\lambda_i$  describe the amount and distribution of that matching. The local systems do not possess definite Schmidt labels before measurement, because the full state is still a coherent superposition. But once the Schmidt-basis measurement is made on one side, the other side is projected into the matching Schmidt vector.

The theorem is therefore the basic structure theorem for pure-state entanglement. It tells us exactly when a pure bipartite state is unentangled, exactly how much local mixedness is caused by entanglement, and exactly which local bases reveal the correlations most simply.

## Product states and entangled states

A pure bipartite state is a product state precisely when its Schmidt rank is one. If

$$r = 1,$$

then the Schmidt decomposition has the form

$$|\psi\rangle_{AB} = |u_1\rangle_A |v_1\rangle_B,$$

which is a product state. Conversely, if  $|\psi\rangle_{AB} = |\alpha\rangle_A |\beta\rangle_B$ , then it already has a one-term Schmidt decomposition.

Thus, for pure states,

$$|\psi\rangle_{AB} \text{ is entangled} \iff r > 1.$$

This criterion is extremely useful because it reduces the question “is this pure state entangled?” to the question “how many nonzero Schmidt coefficients does it have?” Preskill defines the Schmidt number as the number of nonzero terms in the Schmidt decomposition and states that a bipartite pure state is entangled exactly when this number is greater than one.

## Entanglement entropy

Because the Schmidt probabilities are the eigenvalues of both reduced density operators, every entropy of entanglement for a pure bipartite state is a function of the Schmidt probabilities. The most common one is the von Neumann entropy of either reduced state:

$$E(|\psi\rangle_{AB}) = S(\rho_A) = S(\rho_B) = - \sum_i \lambda_i \log \lambda_i.$$

For a product state, the Schmidt spectrum is  $\{1\}$ , so the entropy is zero. For a maximally entangled state of Schmidt rank  $d$ , the Schmidt spectrum is uniform,

$$\lambda_i = \frac{1}{d}, \quad i = 1, \dots, d,$$

so the entropy is

$$E = \log d.$$

This is why the Schmidt decomposition is central in quantum information theory. It turns pure-state entanglement into an ordinary probability distribution. Once we know the numbers  $\lambda_i$ , we know the entanglement entropy, the reduced-state spectra, the Schmidt rank, and the local uncertainty generated by entanglement.

### Example 1: a product state can look like a sum

Consider

$$|\psi\rangle = \frac{|00\rangle + |01\rangle}{\sqrt{2}}.$$

At first sight this is a two-term expression, so a beginner may suspect that it is entangled. But factor it:

$$|\psi\rangle = |0\rangle_A \frac{|0\rangle_B + |1\rangle_B}{\sqrt{2}}.$$

Thus

$$|\psi\rangle = |0\rangle_A |+\rangle_B,$$

where

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

This is a product state. Its Schmidt rank is one, and its only Schmidt probability is  $\lambda_1=1$ . The lesson is important: the number of terms in an arbitrary product-basis expansion is not the Schmidt rank. Entanglement is not detected by counting terms before choosing the correct local bases.

### Example 2: the Bell state

The Bell state

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

is already in Schmidt form. Its Schmidt coefficients are

$$\sqrt{\lambda_1} = \frac{1}{\sqrt{2}}, \quad \sqrt{\lambda_2} = \frac{1}{\sqrt{2}},$$

so

$$\lambda_1 = \lambda_2 = \frac{1}{2}.$$

Tracing out B, we obtain

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

Similarly,

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

The global state  $|\Phi^+\rangle$  is pure, but each local subsystem is maximally mixed. This is one of the clearest operational signatures of entanglement: the whole can be perfectly known while each part alone is maximally uncertain.

The entanglement entropy is

$$E(|\Phi^+\rangle) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1.$$

So a Bell pair contains one ebit of pure bipartite entanglement.

### Example 3: a partially entangled two-qubit state

Consider

$$|\psi\rangle = \sqrt{0.9}|00\rangle + \sqrt{0.1}|11\rangle.$$

This is already in Schmidt form. Its Schmidt probabilities are

$$\lambda_1 = 0.9, \quad \lambda_2 = 0.1.$$

The reduced state of A is

$$\rho_A = 0.9|0\rangle\langle 0| + 0.1|1\rangle\langle 1|.$$

This state is entangled because its Schmidt rank is two. However, it is not maximally entangled because the Schmidt probabilities are not equal. If Alice measures in the Schmidt basis  $\{|0\rangle, |1\rangle\}$ , Bob obtains the same label in his Schmidt basis, but the label 0 occurs much more often than the label 1.

The entanglement entropy is

$$E = -0.9 \log_2(0.9) - 0.1 \log_2(0.1) \approx 0.469.$$

This example shows the difference between being entangled and being maximally entangled. Schmidt rank only tells us that more than one correlated alternative is present. The Schmidt probabilities tell us how evenly the entanglement is distributed.

### Example 4: a state that is maximally entangled in a hidden basis

Consider

$$|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle).$$

This is not obviously in Schmidt form. Write it by grouping the terms according to system A:

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_A \frac{|0\rangle_B + |1\rangle_B}{\sqrt{2}} + \frac{1}{\sqrt{2}}|1\rangle_A \frac{|0\rangle_B - |1\rangle_B}{\sqrt{2}}.$$

Using

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

we get

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_A|+\rangle_B + \frac{1}{\sqrt{2}}|1\rangle_A|-\rangle_B.$$

This is the Schmidt decomposition. The Schmidt probabilities are again

$$\lambda_1 = \lambda_2 = \frac{1}{2}.$$

So  $|\psi\rangle$  is maximally entangled, even though it did not initially look like the standard Bell state. The lesson is that entanglement is invariant under local changes of basis. The Schmidt decomposition finds the local bases in which the correlation structure becomes simple.

### Example 5: unequal dimensions

Let A be a qubit and B be a qutrit. Consider

$$|\psi\rangle = \sqrt{\frac{3}{4}}|0\rangle_A|0\rangle_B + \sqrt{\frac{1}{4}}|1\rangle_A|2\rangle_B.$$

This is already in Schmidt form. The Schmidt rank is two, even though B has dimension three. The reduced states are

$$\rho_A = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|,$$

and

$$\rho_B = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|2\rangle\langle 2|.$$

As operators,  $\rho(A)$  is  $2 \times 2$ , while  $\rho(B)$  is  $3 \times 3$ . Their nonzero eigenvalues are the same, namely  $3/4$  and  $1/4$ . But  $\rho(B)$  has one additional zero eigenvalue because B has an extra unused dimension. This is why the theorem says that the Schmidt rank is at most  $\min(\dim_{\text{mathcal{H}}(A), \dim_{\text{mathcal{H}}(B)})$ .

### How to use the theorem in calculations

In practice, the Schmidt decomposition is usually found in one of two ways.

The first method is the coefficient-matrix method. Choose product bases, write

$$|\psi\rangle = \sum_{a,\mu} M_{a\mu} |a\rangle_A |\mu\rangle_B,$$

and compute the singular value decomposition of  $M$ . The nonzero singular values are the Schmidt coefficients  $\sqrt{\lambda_i}$ , and the left and right singular vectors give the local Schmidt bases. This method is direct and is usually the fastest for explicit finite-dimensional calculations.

The second method is the reduced-density method. Compute

$$\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|)$$

and diagonalize  $\rho_A$ . Its nonzero eigenvalues are the Schmidt probabilities  $\lambda_i$ , and its eigenvectors are the A-side Schmidt vectors. Then compute or infer the matching B-side vectors. This method is physically transparent because it shows immediately how entanglement appears as local mixedness.

For two-qubit states, there is also a useful shortcut. If

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

form the coefficient matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The state is a product state exactly when

$$\det M = ad - bc = 0.$$

If  $ad-bc \neq 0$ , the Schmidt rank is two, so the state is entangled. This determinant test is just the rank test for the coefficient matrix in the special case of two qubits.

## Uniqueness and degeneracy

The Schmidt probabilities  $\lambda_i$  are uniquely determined by the state, up to reordering. The Schmidt vectors are also essentially determined when the nonzero Schmidt probabilities are all distinct. In that case, the only remaining freedom is to multiply  $|u_i\rangle_A$  by a phase and  $|v_i\rangle_B$  by the opposite phase, so that their tensor product is unchanged.

If some Schmidt probabilities are degenerate, the Schmidt bases inside the degenerate subspace are not unique. For example, in the Bell state, the Schmidt spectrum is  $\{1/2, 1/2\}$ , so many different paired local bases can be used to write the same state. This is not a defect. It reflects a physical symmetry: a maximally entangled state has the same amount of correlation in many paired local bases. Preskill notes that when the reduced states have degenerate eigenvalues, the Schmidt decomposition is not fully determined by  $\rho_A$  and  $\rho_B$  alone.

## What the theorem does not say

The Schmidt decomposition is a theorem for pure bipartite states. It does not say that every mixed bipartite state can be written as a single correlated diagonal sum with nonnegative coefficients. Mixed-state entanglement is much more complicated because a mixed state can be decomposed into pure states in many inequivalent ways.

The theorem also does not extend naively to three or more parties. A generic tripartite state

$$|\psi\rangle_{ABC}$$

cannot usually be written as

$$\sum_i \sqrt{\lambda_i} |i\rangle_A |i\rangle_B |i\rangle_C$$

by choosing local orthonormal bases. Bipartite pure-state entanglement has a simple complete structure because matrices have singular value decompositions. Multipartite pure-state entanglement is governed by higher-order tensors, and higher-order tensors do not have an equally simple universal diagonal form.

Finally, the Schmidt decomposition depends on the chosen bipartition. A four-qubit state may have one Schmidt decomposition across the cut  $AB|CD$ , another across  $A|BCD$ , and another across  $AC|BD$ . Entanglement is always entanglement relative to a tensor-product split.

## Final mental image

The Schmidt decomposition says that a pure bipartite state is, after choosing the right local bases, a coherent matching of labels:

$$|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i} |i\rangle_A |i\rangle_B.$$

The labels  $i$  are perfectly correlated. The weights  $\lambda_i$  form the entanglement spectrum. If there is only one label, the state is a product state. If there is more than one label, the state is entangled. If all labels are equally weighted, the state is maximally entangled on its Schmidt support.

This is why the theorem is so useful. It turns the abstract question “what is the structure of a pure entangled state?” into the concrete question “what are the singular values of its coefficient matrix?” It also turns entanglement into local mixedness: the same numbers  $\lambda_i$  are the nonzero eigenvalues of both reduced density operators. In pure bipartite quantum information, almost every calculation begins by finding, using, or implicitly reasoning with this distribution.

## References

Schmidt, Erhard. “Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener.” *Mathematische Annalen* 63, 433–476, 1907.

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 purification.

Preskill, John. *Lecture Notes for Physics 229: Quantum Information and Computation*, Chapter 2, Section 2.4, “Schmidt decomposition.” Available from Caltech.

Watrous, John. *The Theory of Quantum Information*. Cambridge University Press, 2018. See Section 1.1 on Schmidt decompositions and later uses in quantum entropy.

Wilde, Mark M. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017.

Plenio, Martin B., and Shashank Virmani. “An Introduction to Entanglement Measures.” *Quantum Information & Computation* 7, 1–51, 2007.

Miszczak, Jarosław Adam. "Singular Value Decomposition and Matrix Reorderings in Quantum Information Theory." *International Journal of Modern Physics C* 22, 897-918, 2011.

# Document information

## Schmidt Decomposition

---

<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 18:10 KST
<b>Status</b>	Working
<b>Document link</b>	<a href="https://theorytrace.com/projects/mathematical-representation-theorems/documents/un-titled-document/">https://theorytrace.com/projects/mathematical-representation-theorems/documents/un-titled-document/</a>