Module 4 — Entanglement

Quantum Foundations

hbar.university

Tensor Product States

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.1

The composite-system postulate (Node 2.5) declared that the Hilbert space of two systems is the tensor product of their individual Hilbert spaces. This node unpacks the structural consequences. The most important of them — that the tensor product contains vectors which are *not* products — is the entire reason entanglement exists.

Formal Definition

Given two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ with orthonormal bases $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$, the **tensor product** $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is the Hilbert space with orthonormal basis

$$\{|i\rangle_A \otimes |j\rangle_B\}_{i,j}.$$

A general element of $\mathcal{H}_{AB}$ is a complex linear combination

$$|\Psi\rangle_{AB} = \sum_{i,j} c_{ij}\, |i\rangle_A \otimes |j\rangle_B,$$

with coefficients $c_{ij} \in \mathbb{C}$.

A **product state** (or **separable pure state**) is a vector of the form

$$|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\phi\rangle_B$$

for some $|\psi\rangle_A \in \mathcal{H}_A$ and $|\phi\rangle_B \in \mathcal{H}_B$. Equivalently, the coefficient matrix c_{ij} factorizes as $c_{ij} = \alpha_i \beta_j$, i.e., has rank 1.

A vector in $\mathcal{H}_{AB}$ that is *not* a product state is **entangled**.

Intuition

The classical analogue of a composite system is the Cartesian product: if system A can be in 2 states and system B can be in 3, the composite has $2 + 3 = 5$ parameters of state (2 for A, 3 for B). The quantum analogue is the tensor product, which has $2 \cdot 3 = 6$ basis vectors and $6 - 1 = 5$ complex parameters of state — *not* the sum but the *product* of the dimensions.

The "extra" structure that the tensor product provides — beyond what the Cartesian product would — is exactly the space of entangled states. Linear combinations of product states are not in general product states, and there is no classical analogue for this fact.

This is the structural origin of:

- The exponential dimensionality of multi-qubit Hilbert spaces.
- The computational power of quantum simulators.
- The non-local correlations of Bell states.
- The reduction of pure-state information when restricted to subsystems (Module 5).

Structural Role

This node sits between the composite-system postulate (Node 2.5) and everything in the rest of the course that involves more than one subsystem:

- **Separability vs entanglement** (4.2) is built directly on top of this distinction.
- **Schmidt decomposition** (4.3) gives a canonical form for bipartite states.
- **Bell states** (4.4) are the simplest non-trivial entangled vectors.
- **Reduced density operators** (5.2) describe what entanglement looks like from one side only.
- **Quantum information protocols** (Module 7) all operate on tensor product spaces.

Mathematical Development

Tensor product of operators

If $A : \mathcal{H}_A \to \mathcal{H}_A$ and $B : \mathcal{H}_B \to \mathcal{H}_B$ are linear operators, their tensor product $A \otimes B$ acts on $\mathcal{H}_A \otimes \mathcal{H}_B$ by

$$(A \otimes B)(|\psi\rangle_A \otimes |\phi\rangle_B) = (A|\psi\rangle_A) \otimes (B|\phi\rangle_B).$$

This is extended bilinearly to general elements.

Inner product on the tensor product

$$(\langle\psi|_A \otimes \langle\phi|_B)(|\psi'\rangle_A \otimes |\phi'\rangle_B) = \langle\psi|\psi'\rangle_A \cdot \langle\phi|\phi'\rangle_B.$$

This makes $\mathcal{H}_{AB}$ a Hilbert space.

Dimensionality

$$\dim(\mathcal{H}_A \otimes \mathcal{H}_B) = \dim \mathcal{H}_A \cdot \dim \mathcal{H}_B.$$

For n qubits, $\mathcal{H} = (\mathbb{C}^2)^{\otimes n} = \mathbb{C}^{2^n}$. The exponential growth is the structural origin of "quantum advantage."

Worked Example

Two qubits

Let $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$. The four basis vectors of $\mathcal{H}_{AB} = \mathbb{C}^4$ are

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle,$$

where $|ij\rangle$ is shorthand for $|i\rangle_A \otimes |j\rangle_B$.

Product state example:

$$|+\rangle_A \otimes |0\rangle_B = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle).$$

The coefficient matrix is $\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$, which has rank 1. This factors.

Entangled state example (Bell state Φ^+):

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

The coefficient matrix is $\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, which has rank 2. It does not factor. There is no choice of single-qubit states $|\psi\rangle_A, |\phi\rangle_B$ such that $|\Phi^+\rangle = |\psi\rangle_A \otimes |\phi\rangle_B$.

This is the simplest non-trivial entangled state in physics. Every Bell test, every teleportation protocol, and every quantum-key-distribution scheme is built on it.

Cross-Links

- **Linear Algebra:** Tensor products of vector spaces; rank of a matrix.
- **Statistics:** Joint distributions $P(X, Y)$ on a product sample space; independence corresponds to product structure. Entanglement is the quantum analogue of correlation, but it is structurally stronger.
- **Quantum Information:** n-qubit registers, multi-particle states.
- **VQA:** Parameterized circuits on $(\mathbb{C}^2)^{\otimes n}$ explore exactly this exponentially large state space — and barren plateaus arise from precisely this dimensionality (Node 4.4 in VQA course).

Structural Compression

Invariant: A vector in $\mathcal{H}_A \otimes \mathcal{H}_B$ is a product state iff its coefficient matrix has rank 1.

Mechanism: Bilinear extension of basis vector pairing.

Cross-System Echo: Joint distributions in classical statistics are products iff the variables are independent. The tensor-product analogue is sharper: rank-1 vs rank-($>$)1 of the coefficient matrix is a *structural* property, not a statistical one, and it persists regardless of which observables are eventually measured.

Exercises

1. Show that the four states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ form an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$.
2. Express $|+\rangle \otimes |+\rangle$ in the computational basis. Verify it is a product state.
3. Show that the Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ cannot be written as $|\psi\rangle_A \otimes |\phi\rangle_B$ for any $|\psi\rangle, |\phi\rangle \in \mathbb{C}^2$.
4. For the state $|\Psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$, determine whether it is a product state by computing the rank of its coefficient matrix.
5. Show that $\dim((\mathbb{C}^2)^{\otimes n}) = 2^n$ and explain, in one sentence, why this is the source of "quantum advantage."

Separable vs Entangled States

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.2

Node 4.1 distinguished product states from non-product states for pure bipartite systems. For mixed states the question is subtler: a mixed state can be a *statistical mixture* of product states without itself being a product, and any such mixture should still count as classically correlated rather than entangled. The right notion is **separability**, defined by the convex hull of product states. Everything outside that convex hull is entangled, and the boundary between the two — the **separability problem** — is the central problem of finite-dimensional entanglement theory.

Formal Definition

Let $\mathcal{H}_A \otimes \mathcal{H}_B$ be a finite-dimensional bipartite Hilbert space and ρ_{AB} a density operator on it.

Pure-state case. A pure state $|\Psi\rangle_{AB}$ is **separable** (equivalently, a **product state**) if there exist $|\phi\rangle_A \in \mathcal{H}_A$ and $|\chi\rangle_B \in \mathcal{H}_B$ such that

$$|\Psi\rangle_{AB} = |\phi\rangle_A \otimes |\chi\rangle_B.$$

Otherwise it is **entangled**.

Mixed-state case. A density operator ρ_{AB} is **separable** if there exist a probability distribution $\{p_k\}$ and product states $\{\rho_A^{(k)} \otimes \rho_B^{(k)}\}$ such that

$$\rho_{AB} = \sum_k p_k \, \rho_A^{(k)} \otimes \rho_B^{(k)}.$$

Otherwise it is **entangled**. The set of separable states $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is closed and convex; entangled states are everything outside this set.

Without loss of generality, the constituent product states $\rho_A^{(k)}, \rho_B^{(k)}$ can be taken to be pure (any mixed product state can be further decomposed into pure product states). So separability is equivalent to the existence of a representation

$$\rho_{AB} = \sum_k p_k \, |\phi_k\rangle_A\langle\phi_k| \otimes |\chi_k\rangle_B\langle\chi_k|.$$

Intuition

A separable state is one Alice and Bob could prepare from scratch using only **local operations and classical communication** (LOCC). The recipe: a referee picks index k with probability p_k, tells Alice to prepare $|\phi_k\rangle$ and Bob to prepare $|\chi_k\rangle$, and walks away. Whatever ensemble of correlated states this produces is separable, by construction.

The crucial observation is that classical communication between Alice and Bob can produce *correlation* — Alice and Bob's marginal states are correlated through the shared random variable k — but not *entanglement*. Entanglement is the kind of correlation that no LOCC protocol can manufacture from product states. It must be present at the start, distributed by some non-local channel, or generated by a joint quantum interaction.

For pure states the distinction collapses: a pure bipartite state is entangled iff it is non-product, full stop. The mixed-state definition is the natural generalization that respects the convex structure of density operators: any state in the convex hull of products is "as classical as" the products that compose it.

A mixed state can look entangled in one decomposition and separable in another. The defining question is whether *any* convex decomposition into products exists, not whether the most natural one does. This makes the separability problem hard.

Structural Role

Separability is the structural threshold separating quantum-correlated states from classically-correlated ones:

- **Resource theory.** Entanglement is a resource consumed under LOCC. Separable states are "free" — they can be created at no cost — and any monotone of LOCC must be invariant on them. Node 4.5 builds entanglement measures on this skeleton.
- **Entanglement detection.** Telling whether a given ρ_{AB} is separable is, in general, NP-hard (Gurvits 2003). Practical detection uses sufficient conditions: PPT (Peres–Horodecki) for low-dimensional cases, entanglement witnesses (linear functionals separating a state from the convex hull) for the rest.
- **Bell nonlocality (Node 4.4).** Separable states cannot violate any Bell inequality. The converse is false: there exist entangled states with no nonlocal behavior (Werner's example). So entanglement is strictly weaker than Bell nonlocality, but Bell nonlocality is a sufficient certificate for entanglement.
- **Decoherence (Module 5).** A typical effect of decoherence is to drag entangled states across the separability boundary into the separable region — sometimes in finite time ("entanglement sudden death"), unlike the asymptotic decay of coherence in single-system decoherence.

Mathematical Development

Convexity and extreme points

The set $\mathcal{S}$ of separable states on $\mathcal{H}_A \otimes \mathcal{H}_B$ is the convex hull of pure product states $|\phi\rangle\langle\phi| \otimes |\chi\rangle\langle\chi|$. Its extreme points are exactly these pure product states. The full state space is the convex hull of all pure states; $\mathcal{S}$ is a strict, closed convex subset of it (strict whenever $\dim \mathcal{H}_A, \dim \mathcal{H}_B \geq 2$).

By the Carathéodory theorem in dimension $d_A^2 d_B^2 - 1$ (the real dimension of the state space), any separable state can be written as a convex combination of at most $d_A^2 d_B^2$ pure product states. This bounds the search space for explicit separable decompositions but does not make the problem tractable.

The PPT criterion (Peres–Horodecki)

Define the **partial transpose** ρ^{T_B} of ρ_{AB} by transposing only the B-indices in some chosen basis $\{|j\rangle_B\}$:

$$(\rho^{T_B})_{ij,kl} := \rho_{il,kj}.$$

The result depends on the basis, but its spectrum does not.

Theorem (Peres 1996). If ρ_{AB} is separable, then $\rho^{T_B} \succeq 0$.

Proof. If $\rho_{AB} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}$, then $\rho^{T_B} = \sum_k p_k \rho_A^{(k)} \otimes (\rho_B^{(k)})^T$. Each $(\rho_B^{(k)})^T$ is a valid density operator (transposition is positive on Hermitian matrices), so ρ^{T_B} is a convex combination of valid density operators and hence positive. ■

The contrapositive is the **PPT entanglement criterion**: if ρ^{T_B} has a negative eigenvalue, ρ_{AB} is entangled.

Theorem (Horodecki 1996). For $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathbb{C}^2 \otimes \mathbb{C}^3$, PPT is necessary *and* sufficient for separability.

In higher dimensions PPT is only necessary: there exist **PPT entangled states** ("bound entangled" states) that are entangled but cannot be detected by partial transposition. They exhibit a kind of entanglement that cannot be distilled into Bell pairs by LOCC.

Entanglement witnesses

An **entanglement witness** is a Hermitian operator W such that

$$\mathrm{Tr}(W\sigma) \geq 0 \quad \text{for all separable } \sigma, \quad \text{but} \quad \mathrm{Tr}(W\rho) < 0 \text{ for some entangled } \rho.$$

By the Hahn–Banach theorem applied to the closed convex set $\mathcal{S}$, every entangled state is detected by some witness. In practice, witnesses are how entanglement is certified in laboratory experiments — choose a Hermitian W such that $\langle W \rangle$ is measurable with available detectors and outputs a negative value precisely when the state of interest is sufficiently entangled.

LOCC monotonicity

If ρ_{AB} is separable and Λ is any LOCC operation on $\mathcal{H}_A \otimes \mathcal{H}_B$, then $\Lambda(\rho_{AB})$ is separable. LOCC cannot create entanglement from a separable state. This is the structural fact that justifies calling entanglement a *resource* and motivates all of resource-theoretic quantum information.

Worked Example

Werner states

The two-qubit **Werner state** is the one-parameter family

$$\rho_W(p) = p\,|\Phi^+\rangle\langle\Phi^+| + (1-p)\,\frac{I}{4}, \qquad p \in [0,1],$$

where $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. It is the convex combination of a maximally entangled Bell state with the maximally mixed state. The parameter p is the **mixing weight** of the Bell state, also called the *Werner fidelity*.

Compute the partial transpose. In the computational basis, the matrix of $\rho_W(p)$ is

$$\rho_W(p) = \frac{1}{4}\begin{pmatrix} 1+p & 0 & 0 & 2p \\ 0 & 1-p & 0 & 0 \\ 0 & 0 & 1-p & 0 \\ 2p & 0 & 0 & 1+p \end{pmatrix}.$$

The partial transpose with respect to B swaps the off-diagonal coherences $(0,3) \leftrightarrow (1,2)$ (in the basis ordering $00, 01, 10, 11$):

$$\rho_W(p)^{T_B} = \frac{1}{4}\begin{pmatrix} 1+p & 0 & 0 & 0 \\ 0 & 1-p & 2p & 0 \\ 0 & 2p & 1-p & 0 \\ 0 & 0 & 0 & 1+p \end{pmatrix}.$$

The block $\begin{pmatrix} 1-p & 2p \\ 2p & 1-p \end{pmatrix}$ has eigenvalues $(1-p) \pm 2p = 1+p,\ 1-3p$. The other eigenvalues are $1+p$ (twice). So the partial-transpose spectrum is $\{1+p, 1+p, 1+p, 1-3p\}/4$.

Separability boundary. The minimum eigenvalue $(1-3p)/4$ is non-negative iff $p \le 1/3$. By the Horodecki theorem (PPT is sufficient in $\mathbb{C}^2 \otimes \mathbb{C}^2$):

- $p \le 1/3$: $\rho_W(p)$ is separable.
- $p > 1/3$: $\rho_W(p)$ is entangled.

The threshold $p = 1/3$ is the structural "depolarizing limit" of entanglement: any Bell state mixed with more than 2/3 units of white noise becomes classically simulable.

A separable decomposition for $p = 1/3$

At the threshold, an explicit separable decomposition is

$$\rho_W(1/3) = \frac{1}{4}\sum_{i=0}^{3} |\psi_i\rangle\langle\psi_i| \otimes |\psi_i\rangle\langle\psi_i|,$$

where $|\psi_i\rangle$ are four states forming a tetrahedron in the Bloch sphere — for instance, the four corners of an inscribed regular tetrahedron. Verification is a direct computation. This shows that the algebraic threshold from PPT is achieved by an actual LOCC-realizable preparation: pick i uniformly, tell both Alice and Bob to prepare $|\psi_i\rangle$.

A bound entangled example (gestural)

In $\mathbb{C}^3 \otimes \mathbb{C}^3$, the **Horodecki state** built from unextendible product bases gives a state with $\rho^{T_B} \succeq 0$ (PPT) that is nonetheless entangled. No LOCC protocol can extract pure Bell pairs from any number of copies of it. Bound entanglement is the structural sign that entanglement is not a single resource but a hierarchy.

Cross-Links

- **Statistics:** A separable state is the quantum analogue of a *mixture model over independent factor distributions*: a hidden categorical variable k selects a product factorization, and the marginals on A and B are coupled only through k.
- **Quantum Information:** LOCC is the operational paradigm for distributed quantum protocols; separable states are exactly the LOCC-creatable states.
- **Module 4.4 (CHSH):** Bell-inequality violation requires entanglement, but the converse fails — there are entangled states with local hidden-variable models for projective measurements.
- **Module 5 (Decoherence):** Decoherence drags states across the separability boundary, often in finite time ("entanglement sudden death").
- **Module 7 (Quantum Channels):** A channel Λ is **entanglement-breaking** iff $(\Lambda \otimes I)(\rho)$ is separable for every ρ. Such channels are exactly those with measure-and-prepare structure.

Structural Compression

Invariant: The set of separable states is the convex hull of pure product states. Entangled states are exactly the complement — what cannot be built from products by classical mixing.

Mechanism:

$$\rho_{AB} = \sum_k p_k\, \rho_A^{(k)} \otimes \rho_B^{(k)} \quad \Longleftrightarrow \quad \rho_{AB} \text{ is preparable by LOCC.}$$

Cross-System Echo: Hidden-variable models in classical statistics: a classical correlation between A and B can always be modeled by a shared latent variable k inducing conditional independence. Entanglement is precisely the correlation that *cannot* be modeled this way. Bell's theorem (Module 10) makes this structural distinction empirically testable.

Exercises

1. Verify directly that the maximally mixed two-qubit state $I/4$ is separable by exhibiting an explicit decomposition into pure product states.
2. Show that for any pure bipartite state $|\Psi\rangle$, separability (in the mixed-state sense) is equivalent to being a product state. (Hint: a pure state is an extreme point of the state space.)
3. Compute the partial transpose of the Bell state $|\Phi^+\rangle\langle\Phi^+|$ and verify that it has a negative eigenvalue. What is the value?
4. For the Werner state $\rho_W(p)$, compute $\langle\Phi^+|\rho_W(p)|\Phi^+\rangle$ and explain its relationship to the singlet fraction.
5. Construct an entanglement witness W for $|\Phi^+\rangle$ of the form $W = \alpha I - |\Phi^+\rangle\langle\Phi^+|$. Find the smallest α such that $\mathrm{Tr}(W\sigma) \geq 0$ for all separable σ. (Hint: the answer is the maximum overlap of $|\Phi^+\rangle$ with any product state, which is 1/2.)
6. Show that the partial transpose of any product state $\rho_A \otimes \rho_B$ is positive semidefinite. Conclude that the PPT condition is necessary for separability without invoking the convex-combination argument.
7. Argue, structurally, why "entangled" should not be defined as "non-product" for mixed states. What goes wrong if you adopt that definition?

Schmidt Decomposition

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.3

The Schmidt decomposition is the canonical form of a bipartite pure state. It says that any $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, no matter how complicated its expression in a generic product basis, can be brought into a diagonal sum

$$|\Psi\rangle = \sum_i \lambda_i |i_A\rangle \otimes |i_B\rangle$$

by suitable local choices of orthonormal bases. Structurally, this is the singular value decomposition (SVD) wearing physics notation. Operationally, it is the single most useful tool for analysing the entanglement of bipartite pure states: every entanglement quantity that depends only on local degrees of freedom is a function of the **Schmidt coefficients** $\{\lambda_i\}$.

Formal Statement

Theorem (Schmidt decomposition). Let $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure bipartite state, with $\dim \mathcal{H}_A = d_A$ and $\dim \mathcal{H}_B = d_B$. Then there exist:

- orthonormal vectors $\{|i_A\rangle\}_{i=1}^r \subset \mathcal{H}_A$,
- orthonormal vectors $\{|i_B\rangle\}_{i=1}^r \subset \mathcal{H}_B$,
- strictly positive real numbers $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$,

such that

$$|\Psi\rangle = \sum_{i=1}^r \lambda_i \, |i_A\rangle \otimes |i_B\rangle, \qquad \sum_{i=1}^r \lambda_i^2 = 1.$$

The integer $r \leq \min(d_A, d_B)$ is the **Schmidt rank** of $|\Psi\rangle$. The numbers λ_i are the **Schmidt coefficients**, and the sets $\{|i_A\rangle\}, \{|i_B\rangle\}$ are the **Schmidt bases** for the respective sides.

The decomposition is unique up to:

- reordering of equal-magnitude Schmidt coefficients,
- relative phases between paired $(|i_A\rangle, |i_B\rangle)$ compensating each other.

Corollary. $|\Psi\rangle$ is a product state $\iff$ Schmidt rank $r = 1$. In that case the single Schmidt coefficient is $\lambda_1 = 1$.

Corollary. A bipartite pure state is **maximally entangled** when all Schmidt coefficients are equal: $\lambda_i = 1/\sqrt{r}$ for $i = 1, \ldots, r$, with $r = \min(d_A, d_B)$.

Intuition

A generic bipartite pure state, written in a product basis $\{|i\rangle_A \otimes |j\rangle_B\}$, has a coefficient matrix C_{ij} with no particular structure — it can have many nonzero entries, and there is no obvious way to read off "how entangled" the state is from its coefficients.

The Schmidt decomposition is the statement that you can always *rotate the bases on each side independently* — that is, perform local unitaries $U_A \otimes U_B$ — to bring C into diagonal form. The number of nonzero diagonal entries (the rank of C) is then a basis-invariant feature of the state, called the Schmidt rank, and the diagonal entries themselves capture everything there is to say about local entanglement structure.

The diagonal form is exactly the SVD of the coefficient matrix $C = U\Lambda V^\dagger$, with U and V the local unitary changes of basis on A and B, and Λ the diagonal matrix of singular values.

The key intuition: **entanglement is a property of how a bipartite state distributes its amplitude across product directions**, and the Schmidt decomposition is the canonical view in which that distribution is laid out cleanly. A state with one nonzero Schmidt coefficient is a product. A state with all Schmidt coefficients equal is maximally entangled. Everything in between interpolates monotonically.

Structural Role

The Schmidt decomposition organizes essentially every other concept in pure-state entanglement theory:

- **Schmidt rank as entanglement classifier.** Schmidt rank 1 = product = separable. Schmidt rank > 1 = entangled. The maximum possible Schmidt rank is $\min(d_A, d_B)$, achieved by maximally entangled states.
- **Reduced density operators are Schmidt-diagonal.** Tracing out B gives $\rho_A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$, with eigenvalues equal to the squared Schmidt coefficients. The reduced state on A is automatically diagonal in the Schmidt basis. The same holds for ρ_B.
- **Entanglement entropy (Node 4.5).** $S(\rho_A) = -\sum_i \lambda_i^2 \log \lambda_i^2$. This is the canonical entanglement measure for pure states, and it is *symmetric* in A and B: $S(\rho_A) = S(\rho_B)$, since both reduced states share the same Schmidt spectrum.
- **Local unitary equivalence.** Two pure bipartite states are related by a local unitary $U_A \otimes U_B$ if and only if they have the same Schmidt coefficients. The Schmidt spectrum is a complete invariant for the local unitary orbit of pure bipartite states.
- **Purification (Node 5.2).** Every mixed state ρ_A on $\mathcal{H}_A$ is the reduced state of a pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$ (its "purification"). Schmidt decomposition makes this construction immediate: take $|\Psi\rangle = \sum_i \sqrt{p_i}|i\rangle_A \otimes |i\rangle_B$ where $\rho_A = \sum_i p_i |i\rangle\langle i|$.

Mathematical Development

Proof via SVD

Choose any orthonormal product basis $\{|i\rangle_A \otimes |j\rangle_B\}$ of $\mathcal{H}_A \otimes \mathcal{H}_B$. Write

$$|\Psi\rangle = \sum_{i=1}^{d_A}\sum_{j=1}^{d_B} C_{ij}\, |i\rangle_A \otimes |j\rangle_B.$$

The coefficients C_{ij} form a complex $d_A \times d_B$ matrix C. Apply the singular value decomposition:

$$C = U\Sigma V^\dagger,$$

where $U \in U(d_A)$, $V \in U(d_B)$, and Σ is a $d_A \times d_B$ "diagonal" matrix with non-negative real entries σ_i on the diagonal (zero elsewhere). The number of nonzero σ_i is $r := \operatorname{rank} C$.

Substitute back:

$$|\Psi\rangle = \sum_{i,j}\sum_k U_{ik}\sigma_k (V^\dagger)_{kj}\, |i\rangle_A \otimes |j\rangle_B = \sum_{k=1}^{r} \sigma_k \left(\sum_i U_{ik}|i\rangle_A\right) \otimes \left(\sum_j V^*_{jk}|j\rangle_B\right).$$

Define

$$|k_A\rangle := \sum_i U_{ik}|i\rangle_A, \qquad |k_B\rangle := \sum_j V^*_{jk}|j\rangle_B, \qquad \lambda_k := \sigma_k.$$

Then

$$|\Psi\rangle = \sum_{k=1}^{r} \lambda_k\, |k_A\rangle \otimes |k_B\rangle.$$

Since U and V are unitary, $\{|k_A\rangle\}_k$ and $\{|k_B\rangle\}_k$ are orthonormal. Normalization $\langle\Psi|\Psi\rangle = 1$ is equivalent to $\sum_k \lambda_k^2 = 1$ (sum of squared singular values = Frobenius norm squared = 1). ■

Schmidt rank is basis-independent

The Schmidt rank is the rank of any matrix representation of the coefficients, which is invariant under left and right unitary multiplication. So replacing the original product basis by any other orthonormal product basis leaves the Schmidt rank unchanged. This is the structural reason Schmidt rank is a meaningful entanglement classifier.

Reduced density operators

Trace out B from $|\Psi\rangle\langle\Psi|$:

$$\rho_A = \mathrm{Tr}_B\left(\sum_{i,j} \lambda_i\lambda_j |i_A\rangle\langle j_A| \otimes |i_B\rangle\langle j_B|\right) = \sum_{i,j} \lambda_i\lambda_j\, |i_A\rangle\langle j_A|\, \langle j_B|i_B\rangle = \sum_i \lambda_i^2\, |i_A\rangle\langle i_A|,$$

using orthonormality of $\{|i_B\rangle\}$. The reduced state is diagonal in the Schmidt basis with eigenvalues λ_i^2. By the same calculation $\rho_B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$, so

$$\mathrm{spec}(\rho_A) = \mathrm{spec}(\rho_B) \quad \text{(up to padding with zeros)}.$$

This is a structural fact peculiar to *pure* bipartite states: the two marginals have identical nonzero spectra. It fails badly for tripartite or mixed states.

Local unitary equivalence

If $|\Psi\rangle$ and $|\Psi'\rangle$ have the same Schmidt coefficients, they differ by a local unitary: pick U_A sending the Schmidt basis of $|\Psi\rangle$ on A to that of $|\Psi'\rangle$, and similarly U_B. Then $(U_A \otimes U_B)|\Psi\rangle = |\Psi'\rangle$. Conversely, local unitaries preserve Schmidt coefficients, since $U_A \otimes U_B$ acts on the coefficient matrix as $C \mapsto U_A C U_B^T$, and SVD is invariant under left/right unitary multiplication.

The maximally entangled state

A state of the form

$$|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle_A \otimes |i\rangle_B, \qquad d = \min(d_A, d_B),$$

has all Schmidt coefficients equal to $1/\sqrt{d}$, Schmidt rank d, and reduced state $\rho_A = I_d/d$ (maximally mixed). It is the unique (up to local unitaries) state with this property and is the canonical “Bell pair” generalization to dimension d.

Worked Example

Schmidt decomposition of the Bell states

The four Bell states on two qubits:

$$|\Phi^{\pm}\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \qquad |\Psi^{\pm}\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle).$$

For $|\Phi^+\rangle$ the coefficient matrix in the computational product basis is $C = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, already diagonal with singular values $(1/\sqrt{2}, 1/\sqrt{2})$. The Schmidt decomposition is the same as the original expression: Schmidt rank 2, equal Schmidt coefficients $1/\sqrt{2}$, maximally entangled. The reduced state on either side is $\rho_A = \rho_B = I/2$.

For $|\Phi^-\rangle$, $C = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. The SVD is $C = U\Sigma V^\dagger$ with $U = I$, $\Sigma = \frac{1}{\sqrt{2}}I$, $V^\dagger = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. The Schmidt bases are: $|1_A\rangle = |0\rangle, |2_A\rangle = |1\rangle$, $|1_B\rangle = |0\rangle, |2_B\rangle = -|1\rangle$. The decomposition becomes

$$|\Phi^-\rangle = \tfrac{1}{\sqrt{2}}\big(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes (-|1\rangle_B)\big),$$

with both Schmidt coefficients $1/\sqrt{2}$. The minus sign has been absorbed into the local basis on B — it disappears from the *coefficients*, illustrating that the Schmidt spectrum is local-unitary invariant.

For $|\Psi^+\rangle$ the coefficient matrix is $C = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$, the Pauli X matrix scaled by $1/\sqrt{2}$. Its SVD has singular values $(1/\sqrt{2}, 1/\sqrt{2})$. Schmidt rank 2, equal coefficients, maximally entangled.

Conclusion. All four Bell states are maximally entangled with identical Schmidt spectra $(1/\sqrt{2}, 1/\sqrt{2})$. They differ only by local unitaries on Bob's side: $|\Psi^+\rangle = (I \otimes X)|\Phi^+\rangle$, $|\Psi^-\rangle = (I \otimes XZ)|\Phi^+\rangle$, $|\Phi^-\rangle = (I \otimes Z)|\Phi^+\rangle$. The Schmidt decomposition cannot distinguish them, because as far as local entanglement structure goes they are the same state.

A non-maximally entangled example

Consider $|\Psi\rangle = \sqrt{0.9}\,|00\rangle + \sqrt{0.1}\,|11\rangle$. The coefficient matrix is $\mathrm{diag}(\sqrt{0.9}, \sqrt{0.1})$, already in Schmidt form with coefficients $(\sqrt{0.9}, \sqrt{0.1})$. The reduced state is $\rho_A = 0.9|0\rangle\langle 0| + 0.1|1\rangle\langle 1|$, a strongly biased mixture. Entanglement entropy:

$$S(\rho_A) = -0.9\log_2 0.9 - 0.1\log_2 0.1 \approx 0.469 \text{ bits}.$$

Compare with the Bell state's 1 bit. The state is "less entangled" in the precise sense that its Schmidt spectrum is more concentrated.

A non-trivial example requiring the SVD

$$|\Psi\rangle = \tfrac{1}{2}|00\rangle + \tfrac{1}{2}|01\rangle + \tfrac{1}{2}|10\rangle - \tfrac{1}{2}|11\rangle.$$

The coefficient matrix is $C = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = \frac{1}{\sqrt{2}}H$, where H is the Hadamard. The SVD: H is unitary (in fact $H^\dagger H = I$), so the singular values of $H/\sqrt{2}$ are both $1/\sqrt{2}$. The Schmidt decomposition has rank 2 with coefficients $(1/\sqrt{2}, 1/\sqrt{2})$ — this state is also maximally entangled, despite looking quite different from the Bell states. The Schmidt bases on each side are the eigenvectors of $C^\dagger C$ and $CC^\dagger$; for this example they turn out to be $\{|+\rangle, |-\rangle\}$ up to phases. Explicitly:

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}\big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B\big),$$

a Bell state in the Hadamard basis.

Cross-Links

- **Linear Algebra:** The singular value decomposition. Schmidt decomposition = SVD with the matrix indices interpreted as bipartite labels.
- **Statistics:** Low-rank matrix factorization, principal component analysis. The Schmidt rank is the algebraic rank of the bipartite coefficient matrix; PCA picks out the same singular structure for data matrices.
- **Module 4.5 (Entanglement measures):** Entanglement entropy and all related monotones for pure states are functions of the Schmidt coefficients alone.
- **Module 5.2 (Reduced density operators):** Tracing out one subsystem of a pure bipartite state always yields a state diagonal in the Schmidt basis with Schmidt-squared eigenvalues. Purification reverses this construction.
- **Module 7.4 (Quantum teleportation):** The teleportation protocol exploits the Schmidt structure of Bell states — every qubit state can be moved across by local Bell-basis operations.

Structural Compression

Invariant: Every bipartite pure state has a unique-up-to-rotation list of Schmidt coefficients $\{\lambda_i\}$, which is its complete local-unitary invariant.

Mechanism: SVD of the coefficient matrix C_{ij} of $|\Psi\rangle$ in any product basis. The singular values are the Schmidt coefficients; the left/right singular vectors define the Schmidt bases on A and B.

Cross-System Echo: The same algebraic operation appears as PCA in statistics, low-rank approximation in numerical analysis, and latent factor models in machine learning. In quantum mechanics it acquires a structural interpretation as the canonical form of bipartite entanglement: the singular spectrum *is* the entanglement spectrum.

Exercises

1. Compute the Schmidt decomposition of $|\Psi\rangle = \frac{1}{\sqrt{3}}|00\rangle + \frac{1}{\sqrt{3}}|01\rangle + \frac{1}{\sqrt{3}}|10\rangle$. What is its Schmidt rank?
2. Show that for a pure bipartite state, the Schmidt rank equals the rank of ρ_A (and equally of ρ_B).
3. Verify that all four Bell states have the same Schmidt spectrum $(1/\sqrt{2}, 1/\sqrt{2})$, and find local unitaries connecting them to $|\Phi^+\rangle$.
4. Prove that a pure bipartite state $|\Psi\rangle$ is a product state if and only if its reduced density operator ρ_A is pure.
5. Given a mixed state $\rho_A = \sum_i p_i |i\rangle\langle i|$ on $\mathcal{H}_A$, construct an explicit purification $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\rho_A = \mathrm{Tr}_B\, |\Psi\rangle\langle\Psi|$. What is the minimum dimension of $\mathcal{H}_B$?
6. For $|\Psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$, show that the maximum value of the entanglement entropy $S(\rho_A)$ is $\log d$, achieved exactly by maximally entangled states.
7. Argue, structurally, why the Schmidt decomposition cannot be straightforwardly generalized to tripartite pure states. (Hint: try to diagonalize a three-index tensor by local unitaries and see what goes wrong.)

Bell States and the CHSH Inequality

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.4

The CHSH (Clauser–Horne–Shimony–Holt) inequality is the operational form of Bell's theorem. It is the moment in the course where entanglement stops being a piece of mathematical structure and starts producing experimentally falsifiable consequences. The inequality is a tight bound on what *any* local hidden variable theory can predict for a particular sum of correlations between two distant measurement parties; the quantum prediction for a Bell state with the right measurement choices exceeds this bound by a factor of $\sqrt{2}$. The CHSH violation is the simplest, sharpest experimental certificate that the world is not classically correlated.

Formal Statement

The Bell basis

The four **Bell states** on two qubits form an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$|\Phi^+\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \qquad |\Phi^-\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle),$$

$$|\Psi^+\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \qquad |\Psi^-\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

Each is maximally entangled (Schmidt coefficients $(1/\sqrt{2}, 1/\sqrt{2})$, Node 4.3); each pair differs from $|\Phi^+\rangle$ by a single Pauli operator on Bob's qubit. The state $|\Psi^-\rangle$ is the **singlet**, the unique two-qubit state invariant under $U \otimes U$ for every $U \in SU(2)$.

The CHSH operator

Let Alice and Bob each hold one qubit of a shared two-qubit state. Each chooses one of two possible measurements, each with outcomes ± 1:

- Alice measures observable A_0 or A_1, each with eigenvalues ± 1.
- Bob measures observable B_0 or B_1, each with eigenvalues ± 1.

The **CHSH operator** is

$$\mathcal{B} := A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1.$$

For a shared state ρ, the CHSH expectation is $\langle \mathcal{B} \rangle_\rho = \mathrm{Tr}(\mathcal{B}\rho)$.

The classical (local hidden variable) bound

Theorem (CHSH inequality). If the joint distribution of (A_0, A_1, B_0, B_1) admits a local hidden variable model — i.e., there is a probability distribution $p(\lambda)$ and deterministic response functions $a_i(\lambda) \in \{-1, +1\}$, $b_j(\lambda) \in \{-1, +1\}$ such that

$$\langle A_i B_j \rangle = \int p(\lambda)\, a_i(\lambda) b_j(\lambda)\, d\lambda$$

— then

$$|\langle \mathcal{B} \rangle_{\mathrm{LHV}}| \leq 2.$$

The quantum (Tsirelson) bound

Theorem (Tsirelson 1980). For *any* quantum state ρ on a bipartite Hilbert space and *any* observables A_i, B_j with eigenvalues in $[-1, +1]$,

$$|\langle \mathcal{B} \rangle_{\mathrm{QM}}| \leq 2\sqrt{2}.$$

Both bounds are tight: there exist explicit local-deterministic strategies achieving $\langle \mathcal{B} \rangle = 2$, and there is a quantum state and choice of observables achieving $\langle \mathcal{B} \rangle = 2\sqrt{2}$. The latter requires entanglement.

The gap between 2 and $2\sqrt{2}$ is the structural distance between local classical realism and quantum mechanics.

Intuition

The CHSH inequality is the simplest **statement of correlation that any local-realist theory must satisfy**. "Local realist" here means: every measurement outcome is determined by a pre-existing variable λ (realism) that is shared between the parties without any influence travelling faster than light (locality). Under those two assumptions, the algebraic identity

$$A_0(B_0 + B_1) + A_1(B_0 - B_1) = \pm 2$$

holds for any individual run, because exactly one of $B_0 + B_1$ and $B_0 - B_1$ vanishes (the other is ± 2). Averaging over λ gives the bound 2.

Quantum mechanics violates this bound because the four observables A_0, A_1, B_0, B_1 cannot all be assigned simultaneous values. The product $A_0 B_0$ is the expectation of a *jointly measurable* pair (Alice and Bob measure it at spacelike separation), but $A_0 B_0$ and $A_1 B_0$ cannot be jointly measured by Alice (her two observables don't commute), so the algebraic step "factor out B_0" has no quantum analogue. The non-commutation of Alice's observables is what gives quantum mechanics access to the extra $\sqrt{2}$.

The intuitive content is: **non-commuting measurements on an entangled state produce correlations stronger than any pre-existing classical assignment of values can match.** The Bell test is a quantitative way to *measure* this incompatibility from the outside, without ever having to commit to a particular interpretation of the wavefunction.

Structural Role

CHSH is the foundational empirical statement of quantum nonlocality and the starting point for several large structural pieces of quantum theory:

- **Bell's theorem (Module 10).** The CHSH inequality is one operational form of Bell's general result that no local hidden variable theory can reproduce all the predictions of quantum mechanics. Loophole-free experimental violations (Hensen et al. 2015 and successors) established this empirically.
- **Device-independent cryptography.** Because CHSH violation can be observed without trusting the internal workings of measurement devices, it is the structural foundation of device-independent QKD: a protocol whose security depends only on the observed CHSH score, not on assumptions about the apparatus.
- **Tsirelson bound and quantum information geometry.** The Tsirelson bound $2\sqrt{2}$ singles out the *quantum* correlations from the *non-signalling* correlations more generally; the latter could in principle reach 4 (the algebraic maximum). Why nature stops at $2\sqrt{2}$ and not higher is an open foundational question (Popescu–Rohrlich boxes are the standard non-quantum reference).

- **Entanglement certification.** A CHSH score > 2 is a sufficient (though not necessary) certificate of entanglement. Werner showed entangled states with no CHSH violation exist, so CHSH is strictly weaker than entanglement, but it remains the most experimentally accessible test.
- **Self-testing.** A maximal CHSH violation $2\sqrt{2}$ certifies the state and measurements *uniquely*, up to local isometry: it must be a Bell state with the canonical anti-commuting measurements. This rigidity is the basis of self-testing protocols.

Mathematical Development

Derivation of the classical bound

Let λ be the hidden variable and $a_i(\lambda), b_j(\lambda) \in \{-1, +1\}$ the local deterministic response functions. The CHSH expression for a single λ is

$$S(\lambda) := a_0(\lambda)b_0(\lambda) + a_0(\lambda)b_1(\lambda) + a_1(\lambda)b_0(\lambda) - a_1(\lambda)b_1(\lambda).$$

Factor:

$$S(\lambda) = a_0(\lambda)[b_0(\lambda) + b_1(\lambda)] + a_1(\lambda)[b_0(\lambda) - b_1(\lambda)].$$

Since $b_0(\lambda), b_1(\lambda) \in \{-1, +1\}$, one of $b_0 + b_1$ and $b_0 - b_1$ is ± 2 and the other is 0. Therefore $|S(\lambda)| = 2$ for every λ, and

$$|\langle \mathcal{B} \rangle_{\text{LHV}}| = \left| \int p(\lambda) S(\lambda)\, d\lambda \right| \leq \int p(\lambda) |S(\lambda)|\, d\lambda = 2.$$

This is a *deterministic* bound at the level of each λ, so any randomization over λ (or any stochastic LHV that is convex in deterministic LHVs) inherits it. ■

The Tsirelson bound

The CHSH operator can be rewritten as

$$\mathcal{B} = A_0 \otimes (B_0 + B_1) + A_1 \otimes (B_0 - B_1).$$

Compute $\mathcal{B}^2$:

$$\mathcal{B}^2 = A_0^2 \otimes (B_0 + B_1)^2 + A_1^2 \otimes (B_0 - B_1)^2 + \{A_0, A_1\} \otimes (B_0 + B_1)(B_0 - B_1).$$

For observables with $A_i^2 = I, B_j^2 = I$ (eigenvalues ± 1),

$$(B_0 + B_1)^2 + (B_0 - B_1)^2 = 4I, \qquad (B_0 + B_1)(B_0 - B_1) = [B_1, B_0],$$

and similarly the cross term involves $[A_0, A_1] = -[A_1, A_0]$. Combining:

$$\mathcal{B}^2 = 4\, I \otimes I + [A_0, A_1] \otimes [B_0, B_1].$$

Using $\|[A_0, A_1]\| \leq 2$ and $\|[B_0, B_1]\| \leq 2$ (operator norm of a commutator of bounded-by-1 observables),

$$\|\mathcal{B}^2\| \leq 4 + 4 = 8, \qquad \|\mathcal{B}\| \leq 2\sqrt{2}.$$

Hence $|\langle \mathcal{B} \rangle| \leq \|\mathcal{B}\| \leq 2\sqrt{2}$ for any state. ■

Saturating the Tsirelson bound

The bound is saturated by choosing $[A_0, A_1]$ and $[B_0, B_1]$ maximally non-commuting and the state in the joint maximal-eigenvalue eigenspace of $\mathcal{B}$. Concretely, on $|\Phi^+\rangle$ the choice

$$A_0 = Z, \quad A_1 = X, \qquad B_0 = \frac{Z+X}{\sqrt{2}}, \quad B_1 = \frac{Z-X}{\sqrt{2}}$$

gives

$$\langle A_0 B_0\rangle = \frac{1}{\sqrt{2}}, \quad \langle A_0 B_1\rangle = \frac{1}{\sqrt{2}}, \quad \langle A_1 B_0\rangle = \frac{1}{\sqrt{2}}, \quad \langle A_1 B_1\rangle = -\frac{1}{\sqrt{2}},$$

and the CHSH expectation is $4 \cdot \frac{1}{\sqrt{2}} = 2\sqrt{2}$.

Loopholes and loophole-free experiments

Early Bell tests had three structural loopholes: **detection** (low photon counters could allow biased subsamples to fake nonlocality), **locality** (measurements not spacelike separated), and **freedom-of-choice** (the choice of measurement settings could in principle be correlated with the hidden variable). Loophole-free experiments closing all three simultaneously were achieved in 2015 (Hensen et al. with NV centres; Giustina et al. and Shalm et al. with photons). Each used a different physical platform and converged on the same conclusion: the CHSH violation is real, statistically significant, and not an artefact of any classical loophole.

Rigidity (self-testing)

If a black-box experiment achieves CHSH score arbitrarily close to $2\sqrt{2}$, then up to local isometry the shared state must be a Bell state and the measurements must be the maximally non-commuting choices above. This is a **rigidity** statement: the maximal violation pins down the underlying physical implementation, even with adversarial assumptions about the devices. Self-testing protocols use this rigidity to certify quantum hardware in device-independent ways.

Worked Example

Optimal CHSH measurements on $|\Phi^+\rangle$

Take the state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and the measurements

$$A_0 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$B_0 = \frac{Z+X}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad B_1 = \frac{Z-X}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}.$$

Each is Hermitian with eigenvalues ± 1 (so they are valid ± 1-valued observables).

Compute $\langle A_0 \otimes B_0\rangle$ on $|\Phi^+\rangle$. Using the identity $\langle \Phi^+|M \otimes N|\Phi^+\rangle = \frac{1}{2}\operatorname{Tr}(M^T N)$ (exercise),

$$\langle Z \otimes B_0\rangle = \tfrac{1}{2}\operatorname{Tr}(ZB_0) = \tfrac{1}{2} \cdot \tfrac{1}{\sqrt{2}} \operatorname{Tr}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \tfrac{1}{2} \cdot \tfrac{1}{\sqrt{2}} \cdot 2 = \tfrac{1}{\sqrt{2}}.$$

By similar computations:

$$\langle Z \otimes B_1 \rangle = \tfrac{1}{\sqrt{2}}, \quad \langle X \otimes B_0 \rangle = \tfrac{1}{\sqrt{2}}, \quad \langle X \otimes B_1 \rangle = -\tfrac{1}{\sqrt{2}}.$$

CHSH expectation:

$$\langle \mathcal{B} \rangle = \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}} - \left(-\tfrac{1}{\sqrt{2}}\right) = \frac{4}{\sqrt{2}} = 2\sqrt{2}.$$

The Tsirelson bound is saturated. This is the cleanest experimental signature of quantum entanglement.

A separable state cannot violate CHSH

For any separable state $\rho = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}$,

$$\langle \mathcal{B} \rangle_\rho = \sum_k p_k \langle A_0 \rangle_k \langle B_0 \rangle_k + \cdots$$

(each correlator factorizes within a product component). Since each $\langle A_i \rangle_k, \langle B_j \rangle_k \in [-1, 1]$, the same algebraic argument as the LHV bound gives $|\langle \mathcal{B} \rangle_\rho| \leq 2$. So the inequality $\langle \mathcal{B} \rangle > 2$ is a sufficient (though not necessary) certificate of entanglement.

A Werner state's CHSH expectation

For the Werner state $\rho_W(p) = p|\Phi^+\rangle\langle\Phi^+| + (1-p)I/4$, linearity of $\langle \mathcal{B} \rangle$ gives

$$\langle \mathcal{B} \rangle_{\rho_W(p)} = p \cdot 2\sqrt{2} + (1-p) \cdot 0 = 2\sqrt{2}\, p$$

(the maximally mixed state has zero CHSH expectation). The state violates CHSH iff $p > 1/\sqrt{2} \approx 0.707$. Note that for $1/3 < p \leq 1/\sqrt{2}$ the state is entangled (Node 4.2) but does not violate CHSH — there exist entangled states that admit local hidden variable models for projective measurements (Werner 1989). Entanglement and Bell nonlocality are not equivalent.

Cross-Links

- **Statistics:** Conditional independence in causal graphical models. The CHSH inequality is a constraint on the joint distribution of four binary variables under the assumption of a common cause λ screening off all correlations.
- **Quantum Information:** Device-independent quantum key distribution; randomness expansion; self-testing; entanglement certification.
- **Module 4.2 (Separability):** Separable states cannot violate CHSH. The converse (entanglement implies CHSH violation) fails — Werner states between 1/3 and $1/\sqrt{2}$ are entangled but CHSH-local.
- **Module 6 (Interpretations):** Bell's theorem rules out local hidden variable theories but is compatible with non-local hidden variables (Bohmian mechanics), many-worlds (no collapse), and operationalist views (denial of objective values prior to measurement).
- **Module 10.1 (Bell's theorem):** The full statement of Bell's theorem and its experimental history; CHSH is the operational form most often used.

Structural Compression

Invariant:

$$|\langle\mathcal{B}\rangle|_{\text{LHV}} \le 2 \quad < \quad 2\sqrt{2} = |\langle\mathcal{B}\rangle|^{\max}_{\text{QM}} \quad < \quad 4 = |\langle\mathcal{B}\rangle|^{\max}_{\text{NS}}.$$

Mechanism: The CHSH operator on a Bell state has $\mathcal{B}^2 = 4I + [A_0, A_1] \otimes [B_0, B_1]$; maximizing over non-commuting Pauli pairs gives $\|\mathcal{B}\| = 2\sqrt{2}$.

Cross-System Echo: Bell-type inequalities generalize to causal inference in classical statistics: a directed acyclic graph imposes algebraic constraints on observable joint distributions that can be tested empirically. The CHSH inequality is the quantum-violation case, where the *common cause* postulate of classical causal models fails.

Exercises

1. Verify the identity $\langle\Phi^+|M \otimes N|\Phi^+\rangle = \frac{1}{2}\operatorname{Tr}(M^T N)$ for arbitrary 2×2 matrices M, N.
2. Compute $\mathcal{B}^2$ explicitly for the optimal Pauli choice $A_0 = Z, A_1 = X, B_0 = (Z + X)/\sqrt{2}, B_1 = (Z - X)/\sqrt{2}$, and verify that its largest eigenvalue is 8.
3. Show that the singlet $|\Psi^-\rangle$ achieves CHSH violation $-2\sqrt{2}$ with the same measurement choices (note the sign).
4. Derive the CHSH bound for a *stochastic* local hidden variable model (where the response functions a_i, b_j are random rather than deterministic). Show that the bound is the same.
5. Compute $\langle\mathcal{B}\rangle$ for the maximally mixed state $I/4$ with the optimal measurements. Why is the result zero?
6. Construct a local hidden variable model that achieves $\langle\mathcal{B}\rangle = 2$ exactly. (Hint: take λ uniform on a single bit, $a_i(\lambda) = \lambda, b_j(\lambda) = (-1)^{ij}\lambda$ or similar.)
7. Argue, structurally, why the Tsirelson bound $2\sqrt{2}$ and not the algebraic bound 4 is the quantum maximum. What feature of quantum mechanics makes 4 unattainable?

Entanglement Measures

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.5

How much entanglement does a state have? For pure bipartite states the answer is essentially unique up to monotone reparameterization: the **entanglement entropy** of a reduced density operator is the canonical measure. For mixed states the situation fragments — different operational tasks (creating, distilling, certifying entanglement) induce different measures, none of which agrees with any other in general. This node lays out the axiomatic structure of an entanglement measure, the canonical pure-state measure (entanglement entropy), and a survey of the major mixed-state generalizations.

Formal Definition

Axioms of an entanglement measure

An **entanglement measure** is a function $E : \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathbb{R}_{\geq 0}$ satisfying at least:

1. **Vanishing on separable states.** $E(\rho) = 0$ for every separable ρ.
2. **LOCC monotonicity.** $E(\Lambda(\rho)) \leq E(\rho)$ for every LOCC operation Λ. (Often strengthened to *strong monotonicity*: average non-increase under stochastic LOCC.)

Common additional desiderata:

3. **Convexity.** $E(\sum_k p_k \rho_k) \leq \sum_k p_k E(\rho_k)$. Mixing should not increase entanglement.
4. **Continuity.** E is continuous in ρ (in trace norm).
5. **Asymptotic additivity.** $E(\rho^{\otimes n}) = nE(\rho)$ — consistent treatment of n independent copies.
6. **Normalization.** $E(|\Phi^+\rangle\langle\Phi^+|) = \log 2 = 1$ bit for the canonical Bell state.

A function satisfying (1) and (2) is often called an *entanglement monotone*; the stronger desiderata pick out specific physically meaningful measures.

Pure-state entanglement entropy

For a pure bipartite state $|\Psi\rangle_{AB}$ with reduced state $\rho_A := \mathrm{Tr}_B\, |\Psi\rangle\langle\Psi|$, the **entanglement entropy** is

$$E(|\Psi\rangle) := S(\rho_A) = -\,\mathrm{Tr}(\rho_A \log_2 \rho_A) = -\sum_i \lambda_i^2 \log_2 \lambda_i^2,$$

where $\{\lambda_i^2\}$ are the squared Schmidt coefficients of $|\Psi\rangle$ (Node 4.3). By the symmetry of the Schmidt decomposition, $S(\rho_A) = S(\rho_B)$, so the measure is symmetric in $A \leftrightarrow B$.

For pure states this is the **unique** asymptotically continuous, normalized, additive entanglement measure: any other reasonable measure equals it (for pure bipartite states) up to choice of logarithm base.

Mixed-state generalizations

For mixed bipartite states the entropy of the reduced state is *not* a valid entanglement measure — a separable but correlated state has nonzero $S(\rho_A)$ too. The standard mixed-state measures are:

- **Entanglement of formation:** $E_F(\rho) := \min \sum_k p_k S(\rho_A^{(k)})$, minimized over decompositions $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ into pure states. This is a *convex roof* construction.
- **Distillable entanglement:** $E_D(\rho) :=$ the asymptotic rate at which Bell pairs can be extracted from many copies of ρ by LOCC.
- **Relative entropy of entanglement:** $E_R(\rho) := \min_{\sigma \in \mathcal{S}} S(\rho \| \sigma)$, the quantum relative entropy from ρ to the closest separable state.

- **Logarithmic negativity:** $E_N(\rho) := \log_2 \|\rho^{T_B}\|_1$, built from the partial-transpose negativity.

These coincide on pure states ($= S(\rho_A)$) but generically disagree on mixed states.

Intuition

Entanglement is a *resource*, and resource theories ask "how much of it do you have?". The answer depends on what you want to *do* with it. Two natural operational meanings:

- **Cost.** How many Bell pairs must I consume to make a copy of this state by LOCC? → **entanglement of formation**.
- **Yield.** How many Bell pairs can I extract from this state by LOCC? → **distillable entanglement**.

For pure states these two converge to the same number — the entanglement entropy. For mixed states they generically don't, and there are even **bound entangled** states for which the cost is positive but the yield is zero (Node 4.2).

The pure-state case is the "thermodynamic limit" of entanglement theory: in the asymptotic regime of many independent copies, all reasonable entanglement measures collapse to one number, the entropy of the reduced state. The mixed-state case is "non-equilibrium" entanglement theory, where conversion between forms is generally lossy and the resource splinters into a hierarchy.

The structural picture is that the entanglement entropy plays the same role for entanglement that thermodynamic entropy plays for energy: it counts the *exchangeable units* of the resource in the asymptotic regime.

Structural Role

Entanglement measures provide the resource-theoretic backbone of quantum information:

- **Quantification.** They turn the qualitative question "is this state entangled?" into a quantitative one with operational meaning.
- **Conversion.** They bound how much entanglement can be transformed from one form to another by LOCC. The "second law of entanglement" ($E_F \geq E_D$ always) parallels the second law of thermodynamics.
- **Quantum-classical boundary.** Entanglement entropy of subsystems is the order parameter that diagnoses topological and conformal phases of matter (area-law vs volume-law entanglement).
- **Holography (AdS/CFT).** The Ryu–Takayanagi formula relates the entanglement entropy of a boundary region to the area of a minimal surface in the bulk geometry. This is one of the most striking structural appearances of entanglement entropy outside quantum information theory.
- **Decoherence (Module 5).** Open-system dynamics typically transform high-entanglement-entropy pure states into separable mixed states, with the entropy "leaking" into system-environment correlations.
- **VQA.** Trainability of variational quantum circuits depends sensitively on entanglement growth — barren plateaus arise when the circuit's outputs become almost-maximally entangled with auxiliary degrees of freedom.

Mathematical Development

Why the Schmidt spectrum determines pure-state entanglement

Two pure bipartite states are convertible into each other by local unitaries $U_A \otimes U_B$ iff they have the same Schmidt coefficients (Node 4.3). Any entanglement measure invariant under local unitaries (a basic requirement) is therefore a function of the Schmidt spectrum alone. Adding asymptotic additivity and continuity narrows it down to the entropy.

Theorem (Donald–Horodecki–Rudolph 2002, refining Bennett–DiVincenzo–Smolin–Wootters 1996). Among entanglement measures that are normalized, asymptotically continuous, and additive on tensor products of pure bipartite states, the entanglement entropy is unique.

LOCC monotonicity of entropy on pure states

If $|\Psi'\rangle$ is obtained from $|\Psi\rangle$ by LOCC with probability p (for pure-to-pure conversions, this means deterministic conversion via local unitaries plus probabilistic mixing), then by **Nielsen's theorem** (1999),

$$|\Psi\rangle \xrightarrow{\text{LOCC}} |\Psi'\rangle \quad\iff\quad \vec{\lambda}(\Psi) \prec \vec{\lambda}(\Psi'),$$

where $\vec{\lambda}(\Psi) = (\lambda_1^2, \lambda_2^2, \ldots)$ is the squared Schmidt vector and $\prec$ denotes **majorization** ($\vec{\lambda} \prec \vec{\mu}$ iff every cumulative sum of $\vec{\lambda}$ is bounded by the corresponding cumulative sum of $\vec{\mu}$).

The entanglement entropy is Schur-concave (decreases under majorization), so deterministic LOCC $|\Psi\rangle \to |\Psi'\rangle$ requires $S(\rho_A) \geq S(\rho'_A)$. This is the LOCC-monotonicity statement for entropy.

Concentration and dilution

- **Concentration.** n copies of a partially entangled pure state $|\Psi\rangle$ can be converted, asymptotically, into approximately $nS(\rho_A)$ Bell pairs by LOCC. This is *entanglement concentration* (Bennett–Bernstein–Popescu–Schumacher).
- **Dilution.** Conversely, m Bell pairs can be converted into approximately $m/S(\rho_A)$ copies of $|\Psi\rangle$ by LOCC. This is *entanglement dilution*.

The two rates match: pure-state entanglement is asymptotically reversible, and the conversion rate is exactly $S(\rho_A)$.

Entanglement of formation and the convex roof

For a mixed state ρ, define

$$E_F(\rho) := \inf_{\{p_k, |\psi_k\rangle\}} \sum_k p_k S(\mathrm{Tr}_B\, |\psi_k\rangle\langle\psi_k|)\,,$$

the infimum over all pure-state ensembles $\{p_k, |\psi_k\rangle\}$ realizing $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$. The **convex roof** structure ensures vanishing on separable states (which have decompositions into product states with zero Schmidt entropy) and convexity by construction.

For two qubits, Wootters (1998) gave a closed-form expression in terms of the **concurrence** $C(\rho) \in [0, 1]$:

$$E_F(\rho) = h\left(\tfrac{1+\sqrt{1-C(\rho)^2}}{2}\right), \qquad h(x) := -x\log_2 x - (1-x)\log_2(1-x),$$

with $C(\rho) := \max\{0, \sqrt{\mu_1} - \sqrt{\mu_2} - \sqrt{\mu_3} - \sqrt{\mu_4}\}$ and μ_i the eigenvalues (in decreasing order) of $\rho\tilde{\rho}$, where $\tilde{\rho} := (Y \otimes Y)\rho^*(Y \otimes Y)$. This gives the unique closed-form mixed-state entanglement measure for two qubits.

Distillable entanglement and bound entanglement

$E_D(\rho)$ is defined as the supremum, over all LOCC protocols, of the asymptotic rate m/n at which n copies of ρ can be converted into m Bell pairs with vanishing error. It satisfies $E_D(\rho) \leq E_F(\rho)$, with strict inequality for many mixed states. For PPT-entangled states (Node 4.2) the distillable entanglement is exactly zero, even though entanglement of formation is positive — these are the **bound entangled** states.

Logarithmic negativity

The simplest computable mixed-state monotone is the **logarithmic negativity**:

$$E_N(\rho) := \log_2 \|\rho^{T_B}\|_1,$$

where $\|\cdot\|_1$ is the trace norm. This is positive iff the partial transpose has negative eigenvalues, vanishing on PPT states (which include all separable states). It is an entanglement monotone but is *not* convex, and does not coincide with the entropy on pure states. Its main virtue is computability: no minimization or convex roof is required.

Worked Example

Bell state entropy

For $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, the reduced state is

$$\rho_A = \tfrac{1}{2}|0\rangle\langle 0| + \tfrac{1}{2}|1\rangle\langle 1| = \tfrac{I}{2}.$$

Entanglement entropy: $S(\rho_A) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1$ bit.

This is the maximum possible entanglement entropy for a two-qubit state, since ρ_A is maximally mixed in dimension 2 and $\log_2 2 = 1$. All four Bell states share this entropy.

Partially entangled pure state

For $|\Psi\rangle = \cos(\theta/2)|00\rangle + \sin(\theta/2)|11\rangle$, the reduced state is

$$\rho_A = \cos^2(\theta/2)|0\rangle\langle 0| + \sin^2(\theta/2)|1\rangle\langle 1|,$$

with entropy $h(\cos^2(\theta/2))$, the binary entropy. At $\theta = \pi/2$ this equals 1 (Bell state); at $\theta = 0$ it equals 0 (product state $|00\rangle$). The function interpolates monotonically.

Werner state entanglement of formation

For the Werner state $\rho_W(p) = p|\Phi^+\rangle\langle\Phi^+| + (1-p)I/4$ (Node 4.2), Wootters' formula gives

$$C(\rho_W(p)) = \max\{0, (3p-1)/2\}.$$

So the entanglement of formation is

$$E_F(\rho_W(p)) = \begin{cases} 0 & p \le 1/3 \\ h\left(\frac{1+\sqrt{1-((3p-1)/2)^2}}{2}\right) & p > 1/3, \end{cases}$$

continuous in p, zero at the separability boundary, monotonically increasing toward 1 bit at $p = 1$. The structural threshold $p = 1/3$ coincides with the PPT/separability threshold derived in Node 4.2.

Bell state vs. classical correlation

Compare $|\Phi^+\rangle\langle\Phi^+|$ (Bell state, 1 bit of entanglement entropy) with $\frac{1}{2}(|00\rangle\langle 00|+|11\rangle\langle 11|)$ (classically correlated, separable). Both have the same reduced state $\rho_A = I/2$ and so the same *von Neumann entropy of the reduction*. But the second is separable: it has zero entanglement of formation. The entropy of the reduced state is the same, but its *significance* differs. This is why $S(\rho_A)$ is a valid entanglement measure only for *pure* bipartite states.

Cross-Links

- **Statistics:** Mutual information $I(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ generalizes Shannon mutual information; for pure bipartite states it equals $2S(\rho_A)$.
- **Information Theory:** Shannon entropy is the classical analogue of von Neumann entropy; entanglement entropy is the part of quantum mutual information that has no classical realization.
- **Module 4.3 (Schmidt decomposition):** Entanglement entropy is a function of the Schmidt spectrum.
- **Module 4.6 (Monogamy):** Entanglement measures (especially squared concurrence) satisfy monogamy inequalities like CKW.
- **Module 5 (Decoherence):** Decoherence typically reduces entanglement of formation to zero, often in finite time.
- **Module 7 (Quantum Information):** Channel capacities (entanglement-assisted, quantum, private) are formulated as regularized entropy quantities; entanglement measures are the resource currencies.

Structural Compression

Invariant: LOCC monotonicity: any reasonable entanglement measure is non-increasing under local operations and classical communication. Vanishing on separable states is the boundary condition.

Mechanism: For pure states, entanglement entropy $S(\rho_A) = -\sum_i \lambda_i^2 \log \lambda_i^2$. For mixed states, convex-roof extensions (entanglement of formation), regularized rates (distillable entanglement), or norm-based monotones (negativity).

Cross-System Echo: Statistical entropy quantifies *uncertainty* about a system; entanglement entropy quantifies *correlation* with an unobserved partner. The same mathematical object — von Neumann entropy of a reduced density operator — plays both roles, depending on whether the larger system is regarded as a closed pure state or as a mixed ensemble.

Exercises

1. Compute the entanglement entropy of $|\Psi\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)$ on $\mathbb{C}^3 \otimes \mathbb{C}^3$.
2. Show that the entanglement entropy is invariant under local unitaries $U_A \otimes U_B$.
3. For the partially entangled state $|\Psi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$, plot $S(\rho_A)$ as a function of θ. Where is the maximum?
4. Show that for any pure state $|\Psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$, $S(\rho_A) \leq \log_2 d$, with equality iff $|\Psi\rangle$ is maximally entangled.
5. Verify that the classically correlated state $\frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$ is separable but has $S(\rho_A) = 1$ bit. Why doesn't this count as 1 bit of entanglement?
6. Compute the logarithmic negativity $E_N(\rho_W(p))$ of the Werner state and compare its dependence on p with the entanglement of formation.
7. Argue, structurally, why pure-state entanglement is asymptotically reversible (one number suffices) while mixed-state entanglement is not (irreducible hierarchy of measures).

Monogamy of Entanglement

Position in Knowledge Graph

System: Quantum Foundations **Structural Domain:** Entanglement **Node:** 4.6

Quantum entanglement cannot be freely shared. If two parties A and B are maximally entangled, then A cannot be entangled with any third party C at all. This is structurally different from classical correlation, which can be shared arbitrarily: three people can be perfectly correlated on a shared random bit. Monogamy is one of the sharpest structural distinctions between quantum and classical information, and it is the structural reason quantum key distribution is secure — any eavesdropper's entanglement with the transmitted state costs entanglement between the legitimate parties in a quantifiable way.

Formal Statement

The CKW (Coffman–Kundu–Wootters) inequality

Let A, B, C be three qubits in a joint state ρ_{ABC}. Define the **tangle** $\tau(\rho) := C(\rho)^2$, the squared concurrence (Node 4.5). Then

$$\tau(A, B) + \tau(A, C) \leq \tau(A, BC),$$

where $\tau(A, B)$ and $\tau(A, C)$ are the tangles of the two-qubit reduced states $\rho_{AB} = \mathrm{Tr}_C\, \rho_{ABC}$ and $\rho_{AC} = \mathrm{Tr}_B\, \rho_{ABC}$, and $\tau(A, BC)$ is the tangle of the bipartition A versus (B, C) treated as a single four-dimensional subsystem.

The inequality is **tight**: it is saturated by the GHZ state, and its deficit (the difference between right and left sides) is called the **three-tangle** or **residual tangle**,

$$\tau_{ABC} := \tau(A, BC) - \tau(A, B) - \tau(A, C),$$

which is a genuine measure of three-party entanglement that cannot be captured by bipartite tangles alone.

General formulation

For higher dimensions and larger numbers of parties, the CKW inequality has been generalized in various ways. The cleanest version is due to Osborne and Verstraete (2006):

$$\sum_{i=2}^{n} \tau(A_1, A_i) \leq \tau(A_1, A_2 \cdots A_n)$$

for any n-qubit pure state. The structural content is the same: the total bipartite entanglement between one party and each of the others is bounded above by the bipartite entanglement between that party and the rest of the system considered as a whole.

Not every entanglement measure is monogamous in this sense; for instance, entanglement of formation is *not* monogamous on three qubits. Squared concurrence (the tangle) is the canonical choice for which the CKW inequality holds.

Intuition

Classical correlations are sharable without limit. If Alice, Bob, and Charlie each hold a copy of a shared random bit x, Alice is perfectly correlated with Bob *and* perfectly correlated with Charlie, simultaneously. There is no upper bound on how many parties can share the same classical information.

Quantum entanglement cannot be copied this way, structurally. The reason traces back to the **no-cloning theorem** (Node 7.2): if Alice could share a maximally entangled Bell pair with both Bob and Charlie simultaneously, she could (in effect) clone the Bell pair by tracing out one side, violating no-cloning. Monogamy is the quantitative refinement of this impossibility: not just "you can't share Bell pairs perfectly with two partners," but "the more entanglement you have with Bob, the less you can have with Charlie," with an exact trade-off relation.

The paradigmatic fact: **if A is in a pure entangled state with B, then C is completely uncorrelated with A.** Mathematically, if ρ_{AB} is pure, then $\rho_{ABC} = \rho_{AB} \otimes \rho_C$, so C factorizes off. There is no room for A to be entangled with anything else when its entanglement with B is already "saturated."

The CKW inequality makes this precise for partial entanglement: whatever entanglement A has with B subtracts from the entanglement it can have with C, in a specific quadratic way.

Structural Role

Monogamy is one of the load-bearing structural facts of quantum information theory:

- **Quantum key distribution (QKD) security.** In BB84 or E91 protocols, Alice and Bob measure correlated halves of a Bell state. If an eavesdropper Eve has entanglement with the qubits in transit, monogamy forces a reduction in Alice–Bob correlations that Alice and Bob can detect via Bell tests. The security of QKD is literally the CKW inequality applied to Alice–Bob–Eve.
- **No-cloning reinforcement.** Monogamy is the continuous-valued form of no-cloning. No-cloning says you can't copy arbitrary quantum states; monogamy says that even partial copies (approximate cloners) must degrade the original's entanglement with the environment.
- **Constraint on multi-party entanglement.** Monogamy means that entanglement in an n-party system is a constrained resource; a party cannot be simultaneously maximally entangled with all others. This disciplines the structure of multipartite entanglement into hierarchies like GHZ-class vs. W-class.
- **Holography and the ER=EPR conjecture.** In AdS/CFT, the bulk geometry dual to an entangled boundary state has a wormhole structure. Monogamy of entanglement on the boundary corresponds to the topological fact that one side of a wormhole can connect to only one other side — a striking structural alignment between quantum information and spacetime geometry.
- **Many-body physics.** Monogamy constrains the entanglement structure of ground states of local Hamiltonians and is behind the emergence of area laws for entanglement entropy in gapped systems.

Mathematical Development

Proof sketch of CKW for three qubits

Coffman, Kundu, and Wootters (2000) proved the inequality for pure three-qubit states by direct calculation using Wootters' closed-form expression for the concurrence. The key steps:

1. For a pure three-qubit state $|\psi\rangle_{ABC}$, the bipartition $A|BC$ treats BC as a four-dimensional system and $\tau(A, BC) = 4 \det \rho_A$, where $\rho_A = \mathrm{Tr}_{BC} |\psi\rangle\langle\psi|$. For a qubit, $4 \det \rho_A = 1 - |\vec{r}|^2$ where $\vec{r}$ is the Bloch vector.
2. The two-qubit reduced states ρ_{AB}, ρ_{AC} are generally mixed; their concurrences are computed by Wootters' formula.
3. Algebraic manipulation (lengthy but direct) shows $\tau(A, B) + \tau(A, C) \leq 4 \det \rho_A = \tau(A, BC)$, with equality iff the three-tangle vanishes.

The extension to mixed states is by convex roof: the CKW inequality holds for any convex roof extension of the squared concurrence, and equals the pure-state expression on extremal points.

Three-tangle as a genuine tripartite invariant

The residual $\tau_{ABC} := \tau(A, BC) - \tau(A, B) - \tau(A, C)$ measures the entanglement that is *not* captured by any of the bipartite reductions. It is symmetric under permutation of A, B, C (a nontrivial fact), is invariant under local unitaries, and vanishes on states that are in the "W class" (discussed below). It characterizes a genuinely multipartite form of entanglement.

Why the squared concurrence?

Not every entanglement monotone satisfies a monogamy inequality. Entanglement of formation E_F, despite being convex and monotonic under LOCC, is **not monogamous**: there exist three-qubit states with $E_F(A, B) + E_F(A, C) > E_F(A, BC)$. The tangle (squared concurrence) is special because its convex-roof structure interacts correctly with the bipartition. Later generalizations (Bai et al., 2014) showed that *any* α-th power of the concurrence with $\alpha \geq 2$ is monogamous, while $\alpha < 2$ fails.

Structurally, monogamy is a feature of the *geometry* of state space under specific measures, not a universal property of "entanglement." The right way to read the CKW inequality is as picking out τ as the *monogamous* measure, rather than claiming that entanglement in general is monogamous.

The Osborne–Verstraete extension

For pure states of n qubits, Osborne and Verstraete proved:

$$\sum_{i=2}^{n} \tau(A_1, A_i) \leq \tau(A_1, A_2 \cdots A_n) = 4 \det \rho_{A_1}.$$

The right side is bounded by 1 (for a qubit), so the total tangle of A_1 distributed among the remaining $n - 1$ parties is bounded by 1. This is the quantitative form of "you cannot simultaneously maximally entangle one qubit with many others."

Worked Example

GHZ state

The three-qubit GHZ state

$$|\mathrm{GHZ}\rangle = \tfrac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$$

has reduced states

$$\rho_{AB} = \tfrac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|), \qquad \rho_A = \tfrac{1}{2} I.$$

The two-qubit reduced state ρ_{AB} is separable (a classical mixture of $|00\rangle$ and $|11\rangle$), so its concurrence $C(A, B) = 0$ and tangle $\tau(A, B) = 0$. By symmetry $\tau(A, C) = 0$. The bipartite tangle is

$$\tau(A, BC) = 4 \det \rho_A = 4 \cdot \tfrac{1}{4} = 1.$$

CKW inequality:

$$\tau(A,B) + \tau(A,C) = 0 \le 1 = \tau(A,BC),$$

satisfied with maximum deficit. The three-tangle is $\tau_{ABC} = 1$ — the GHZ state has all of its entanglement in the genuinely tripartite mode.

The structural picture: GHZ distributes entanglement so that *no pair* of qubits is entangled when the third is traced out, yet the global state is maximally entangled against any single-qubit cut. This is genuinely "non-bipartite" entanglement, undetectable by any bipartite measure.

W state

The three-qubit W state

$$|W\rangle = \tfrac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$$

has reduced states

$$\rho_{AB} = \tfrac{1}{3}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \rho_A = \tfrac{1}{3}|1\rangle\langle 1| + \tfrac{2}{3}|0\rangle\langle 0|.$$

Wootters' formula gives $C(A,B) = 2/3$, so $\tau(A,B) = 4/9$. By symmetry $\tau(A,C) = 4/9$. Compute

$$\tau(A,BC) = 4\det\rho_A = 4 \cdot \tfrac{2}{3} \cdot \tfrac{1}{3} = \tfrac{8}{9}.$$

CKW inequality:

$$\tau(A,B) + \tau(A,C) = \tfrac{4}{9} + \tfrac{4}{9} = \tfrac{8}{9} = \tau(A,BC).$$

The inequality is **saturated**, and the three-tangle of the W state is *zero*:

$$\tau_{ABC}(W) = 0.$$

Structurally: the W state distributes its entanglement entirely into bipartite correlations — every pair is entangled, but there is no genuinely tripartite residue. This is the **W class** of tripartite entanglement, distinct from the GHZ class: local unitaries cannot convert between the two, even with finite probability.

Taxonomy of three-qubit entanglement. Tripartite pure states come in two structurally different families:

- **GHZ class:** $\tau_{ABC} > 0$. Genuinely tripartite entanglement. Tracing out any party destroys all remaining correlation.
- **W class:** $\tau_{ABC} = 0$. Pairwise entanglement only. Robust to the loss of one party — the remaining pair is still entangled.

This is the structural explanation for why "a Bell pair plus a spectator" (product class), the W state (pairwise), and the GHZ state (tripartite) are the irreducible patterns of three-qubit entanglement.

QKD security sketch

In an E91-style protocol, Alice and Bob share Bell pairs distributed through a channel. An eavesdropper Eve can, in principle, intercept and entangle an ancilla with each pair. The tripartite state is $|\Psi\rangle_{ABE}$, and monogamy bounds:

$$\tau(A, B) + \tau(A, E) \leq \tau(A, BE).$$

If Alice and Bob verify by Bell tests that their two-party tangle is close to 1 (maximal entanglement), then $\tau(A, E)$ is forced close to 0 — Eve cannot be entangled with A. No entanglement with A means Eve cannot extract information about the shared key in a coherent way. This is the **structural origin of QKD security**.

Cross-Links

- **Module 4.5 (Entanglement measures):** The squared concurrence (tangle) is the monogamous measure; not all entanglement measures share this property.
- **Module 7.2 (No-cloning):** Monogamy is the continuous quantitative form of the no-cloning theorem.
- **Quantum Information:** QKD security proofs rely on monogamy to bound an eavesdropper's correlations.
- **Holography:** ER=EPR and Ryu–Takayanagi align monogamy with the topological structure of bulk wormholes.
- **Statistics:** Classical mutual information is not monogamous — perfect correlation can be shared freely. Monogamy is a structural feature with no classical analogue.

Structural Compression

Invariant:

$$\tau(A, B) + \tau(A, C) \leq \tau(A, BC).$$

Bipartite entanglements sharing a common party satisfy a quadratic constraint.

Mechanism: For three-qubit pure states, direct algebraic proof using Wootters' concurrence formula. Generalizations via convex-roof extensions and the Osborne–Verstraete theorem for n qubits.

Cross-System Echo: No classical analogue. Classical correlations and mutual information are freely sharable; entanglement is not. Monogamy is one of the structural features of quantum correlation that most sharply distinguishes it from classical statistics, and is the quantitative backbone of quantum cryptography.

Exercises

1. Verify by direct computation that the two-qubit reduced state ρ_{AB} of the GHZ state is separable, and hence has zero concurrence.
2. Compute the reduced states $\rho_{AB}, \rho_{AC}, \rho_A$ of the W state and verify the concurrence values $C(A, B) = C(A, C) = 2/3$.
3. Show that for any pure tripartite state, $\tau(A, BC) = 4 \det \rho_A$ where ρ_A is the single-qubit reduced state. (Hint: the bipartition is pure, so use Schmidt decomposition and compute directly.)
4. Construct a three-qubit state that saturates the CKW inequality with $\tau(A, B) = \tau(A, C) = \tau(A, BC)/2$. Is it in the GHZ class or the W class?
5. Show that entanglement of formation is *not* monogamous by giving a three-qubit example with $E_F(A, B) + E_F(A, C) > E_F(A, BC)$. (This requires looking up a counterexample from the literature.)

6. Explain, using monogamy, why an eavesdropper Eve cannot perfectly clone a qubit Alice sends to Bob without destroying the Alice–Bob entanglement needed to detect her. Make the bound quantitative.
7. Argue, structurally, why monogamy of entanglement is a direct consequence of the no-cloning theorem, and why its *quantitative* form (CKW) goes beyond no-cloning by bounding partial correlations rather than just exact cloning.