Bell States and the CHSH Inequality

Quantum Foundations · Module 4 — Entanglement

Node 4.4

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

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

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 $[-1, +1]$,

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

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 $|\langle \mathcal{B}\rangle_\rho| \le 2$. 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 $1/3 < p \le 1/\sqrt{2}$ 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:

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

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?

Quantum Foundations · Module 4 — Entanglement · Node 4.4
Concepts: entanglement, hypothesis-testing, conditional-probability
Prerequisites: 4.1, 4.3, 2.3
Enables: 10.1