

Scaling Laws for Agreement-Filtered CHSH Violations in Objectivity-Loophole Tests

Rishabh Venkataramani¹

Received December 27, 2025

Accepted April 13, 2026

Electronic access May 31, 2026

Bell tests are experiments that demonstrate how particles communicate, or “entangled,” despite being far apart. The Wigner’s friend thought experiment builds on this, raising questions about what occurs when the observers themselves are incorporated into the quantum system. Recent experimental implementations of the Wigner’s friend thought experiment have performed the so-called “agreement filter,” a postselection rule requiring unanimous results among redundant observers prior to announcing a result objectively. These demonstrations, including a test in 2025 on IBM’s superconducting quantum hardware, have been restricted to relatively few friends per wing. We present the first closed-form derivation of scaling laws for the CHSH parameter in agreement filtering with N friends per wing. We obtain analytic violation thresholds under different realistic noise models, including measurement errors, depolarization, noise in hybrid photonic-superconducting systems, and, beyond these, correlated errors and the general k -of- N voting rules. Our results determine arrangements under several realistic noise models and find that, under current device parameters, a friend count of $N \approx 3$ optimizes results.

Introduction

One of the central questions in quantum mechanics concerns the objectivity of measurement outcomes: whether different observers must necessarily agree on what has been measured¹. While quantum theory successfully predicts experimental results, it does not clearly specify whether measurement outcomes can be treated as observer-independent facts² when observers themselves are included within the quantum description. This issue is highlighted by the Wigner’s friend thought experiment, in which an observer inside a laboratory performs a measurement and records a definite outcome, while an external observer treats the entire laboratory, including the friend and the measured system, as a single quantum system.

The standard CHSH inequality Bell test consists of two parties sharing states and making measurements^{3,4}. Under classical assumptions, the standard CHSH inequality constrains correlations to $|S| \leq 2$, while quantumly entangled states can get up to $|S| \leq 2\sqrt{2}$, a limit known as the Tsirelson bound⁵. Recent tests on quantum computers have been inspired by the Wigner’s friend thought experiment, where measurements must be agreed upon by all of the observers of a party⁶. In the Wigner’s friend problem, Wigner measures the quantum system and sees a definite result, while “the friend” measures the whole system, including the lab. The “agreement filter” is used for all friends to agree on the measurement.

The CHSH parameter is defined as

$$S = |E(0,0) - E(0,1) + E(1,0) + E(1,1)|$$

The correlators $E(a,b)$ represent the average product of the measurement outcomes for each combination of settings a and b , taking values between -1 and $+1$. In classical, local realistic theories, S is bounded by 2, but quantum mechanics allows stronger correlations that can reach the Tsirelson limit of $2\sqrt{2}$. This violation is a signature of quantum entanglement⁷ and the nonlocal nature of quantum mechanics.

The term “objectivity-loophole” refers to the possibility that measurement outcomes used in a Bell test may not correspond to observer-independent, definite facts. Extended Wigner’s friend scenarios demonstrate that quantum theory allows consistent descriptions in which different observers assign incompatible accounts to the same measurement event⁸. If Bell violations are demonstrated without verifying agreement between multiple observers on each wing, one may question whether the outcomes entering the CHSH correlator represent shared, objective records. The agreement filter utilizes a test of objectivity by requiring N redundant observers per wing to report unanimous outcomes before a trial is accepted. By enforcing inter-observer consensus, the protocol strengthens the requirement of definite, observer-independent outcomes.

In objectivity-loophole tests, this framework is extended to include multiple observers (“friends”) who redundantly record the measurement outcome before an agreement filter determines whether their records are consistent. Only when all friends on each side report the same result is the trial valid. This process connects fundamental questions about the objec-

¹ Vista Del Lago High School, Folsom, CA; Email: rven2308@gmail.com

tivity of measurement⁹ with testable predictions in quantum hardware.

Bell's original inequality established experimentally testable bounds separating local realism from quantum predictions³, and its CHSH formulation provided a practically implementable version of this test^{4,10}. Subsequent experimental demonstrations confirmed violations of these inequalities and established nonlocality as an empirical feature of quantum mechanics¹¹.

Extensions of Bell scenarios to include nested or multiple observers have been developed in the context of the Wigner's friend paradox and local friendliness inequalities. A strong no-go theorem showed that observer-independent facts are incompatible with quantum predictions under reasonable physical assumptions¹², and related analyses further clarified the constraints imposed on classical descriptions of measurement outcomes^{13,14}.

Recent experimental efforts have begun implementing simplified versions of these extended scenarios. Proietti *et al.* reported an experimental test of local observer independence using entangled photons¹⁵, while Bednorz *et al.* implemented objectivity-loophole tests on publicly accessible quantum hardware⁶. These demonstrations primarily focused on configurations with small numbers of friends. The quantitative behavior of agreement-filtered correlations as the number of redundant observers grows has not been systematically analyzed. We derive closed-form scaling laws describing how agreement-filtered CHSH violations behave as N increases, providing analytic expressions for $S(N)$ under independent, depolarizing, and correlated noise that link multi-observer foundational scenarios to experimentally accessible hardware regimes.

Theoretical Background

The Wigner's friend scenario extends the quantum measurement problem to multiple observers¹⁶. In the standard setup, a quantum system is measured by an observer inside a laboratory (the "friend"), while an external observer (Wigner) treats the combined system of the friend and lab as a quantum state. The paradox arises when Wigner's and the friend's descriptions of the same event become inconsistent: one assigns a definite outcome, while the other describes a superposition of possibilities. This conflict motivates a quantitative study of observer agreement and the limits of objectivity in quantum mechanics.

The Objectivity Loophole

In standard Bell tests, several well-known loopholes must be addressed, including the detection loophole (inefficient sampling)^{17,18}, the locality loophole (space-like separation

of measurement events), and the freedom-of-choice loophole (independence of measurement settings)¹⁹. These concern experimental imperfections that could allow local hidden-variable explanations of observed violations²⁰.

The objectivity loophole is conceptually distinct. It concerns whether measurement outcomes can be treated as definite, observer-independent facts prior to being jointly compared. In extended Wigner's friend scenarios, different observers may assign incompatible descriptions to the same measurement event. If one assumes that outcomes are objective—i.e., that all observers can consistently assign the same definite result—then Bell-type inequalities can be derived under additional assumptions about observer independence.

Objectivity-loophole tests therefore examine whether experimentally observed correlations remain compatible with assigning joint, observer-independent outcomes across multiple measurement agents. It has been shown that assuming single, observer-independent outcomes for all agents in nested measurement scenarios leads to logical inconsistencies within quantum theory itself, motivating operational tests of observer agreement²¹. In addition, it has been demonstrated that under locality and freedom-of-choice assumptions, jointly assigning definite outcomes to multiple observers yields inequalities that are violated by quantum predictions¹³. Moreover, violations of local friendliness inequalities constrain causal explanations that rely on observer-independent facts, clarifying the foundational implications of such tests¹⁴.

Together, these works motivate one to treat objectivity as an operational constraint rather than an assumption. In this context, the agreement filter explicitly enforces inter-observer consensus before assigning a classical record, thereby closing the objectivity loophole and enabling quantitative tests of how strengthened objectivity requirements suppress nonlocal correlations.

A natural framework for formalizing these ideas is through extensions of the CHSH inequality to "local friendliness" tests, which combine Bell-type nonlocal correlations with nested measurement structures¹². The standard CHSH inequality involves two observers, A and B , who perform local measurements a and b in entangled pairs. Correlators are defined as

$$E(a, b) = \langle A(a)B(b) \rangle,$$

and combined into the CHSH parameter,

$$S = E(a_0, b_0) - E(a_0, b_1) + E(a_1, b_0) + E(a_1, b_1).$$

Classically, $|S| \leq 2$, while quantum mechanics allows violations up to $|S| = 2\sqrt{2}$, the Tsirelson bound.

In the context of Wigner's friend extensions, the "friend" acts as an internal observer who performs a first measurement, while Wigner acts as a higher-level observer on the combined

system. When multiple redundant friends are introduced per wing, a postselection rule known as the *agreement filter* can be applied: only runs in which all friends on a side report the same outcome are counted as valid measurement results. This rule operationalizes the notion of objectivity by requiring unanimous internal agreement before an outcome is registered externally.

The introduction of the agreement filter modifies both the statistics of accepted runs and the effective correlations contributing to S . As a result, the CHSH value becomes a function of the friend count N , the voting rule, and an underlying noise model. Theoretical modeling of these dependencies yields scaling laws for $S(N)$.

Agreement Filter Formalism

Consider a CHSH test with two parties, A and B . Each party chooses one of two binary-outcome settings $a \in \{0, 1\}$ and $b \in \{0, 1\}$ respectively, and each individual measurement yields an outcome in $\{+1, -1\}$. In the agreement-filtered scenario we attach N redundant observers (“friends”) to each wing. Each friend independently records a binary outcome $x_i \in \{+1, -1\}$ for party A (and similarly y_j for party B). The agreement filter converts the N bits on each wing into a single output

$$X, Y \in \{+1, -1, 0\},$$

where $+1$ or -1 is reported if the friends unanimously report $+1$ or -1 respectively, and 0 denotes disagreement, and the trial is discarded.

To formalize the noise model, let $s \in \{+1, -1\}$ denote the latent true sign that would be obtained in the absence of friend errors. For each wing, the agreement filter outputs a reported sign $X \in \{+1, -1\}$ on accepted runs (and 0 otherwise, in which case the run is discarded). All expectations below are taken over the independent error randomness of the N friends, conditioned on acceptance by the agreement filter.

Now suppose each friend independently flips the true bit with probability e (so friends report the correct bit with probability $1 - e$). For a single wing with N friends the probability that all friends agree (either all correct or all flipped) is

$$p_{\text{agree}}(N, e) = (1 - e)^N + e^N,$$

The acceptance probability is plotted as a function of N in Figure 1 for several error rates, illustrating the rapid decay in accepted trials as redundancy increases. The conditional sign fidelity of the reported outcome given acceptance is

$$\eta_N(e) = \mathbb{E}[X/s \mid \text{agree}] = \frac{(1 - e)^N - e^N}{(1 - e)^N + e^N},$$

where $s \in \{+1, -1\}$ denotes the latent true sign. Equivalently,

$$\eta_N(e) = \tanh(N \operatorname{arctanh}(1 - 2e)).$$

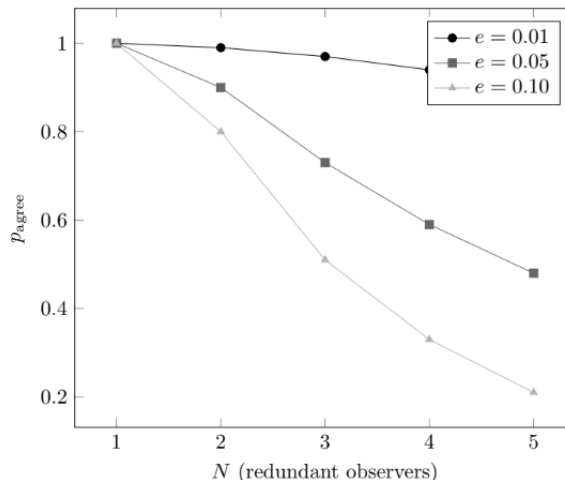


Fig. 1 Unanimous acceptance probability $p_{\text{agree}} = (1 - e)^N + e^N$ for several error rates.

The corresponding fidelity factor is shown in Figure 2, confirming that conditional accuracy improves with N for small error rates.

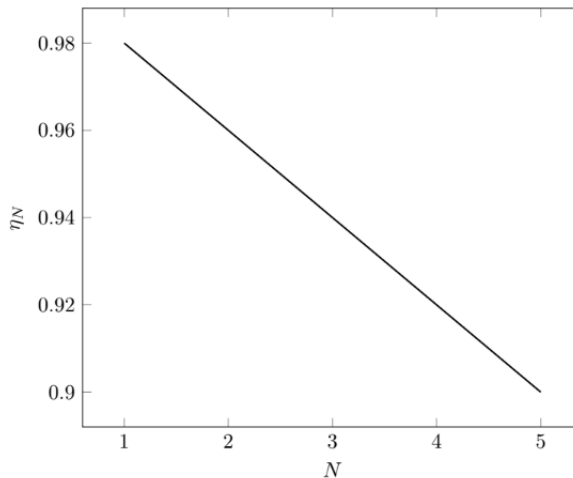


Fig. 2 Fidelity factor $\eta_N = \tanh(N \operatorname{arctanh}(1 - 2e))$ for small independent error rate.

When both wings apply identical filters (possibly with different error rates e_A, e_B), the joint acceptance probability for a given run is

$$\alpha(N, e_A, e_B) = p_{\text{agree}}(N, e_A) p_{\text{agree}}(N, e_B),$$

so only a fraction α of the original trials contribute to postselected correlators.

Integration into CHSH

Let $E_{\text{ideal}}(a, b)$ denote the ideal correlator (no friend errors, no filter). Because the agreement filters on A and B independently multiply the local outcomes by random signs with means $\eta_N(e_A)$ and $\eta_N(e_B)$ respectively, the conditional two-party correlator on accepted runs becomes

$$E_N(a, b) = \eta_N(e_A)\eta_N(e_B)E_{\text{ideal}}(a, b).$$

Consequently, the CHSH parameter after filtering is

$$\begin{aligned} S(N, e_A, e_B) &= \\ &|E_N(0, 0) - E_N(0, 1) + E_N(1, 0) + E_N(1, 1)| \\ &= \eta_N(e_A)\eta_N(e_B)S_{\text{ideal}}, \end{aligned}$$

where S_{ideal} is the CHSH value in the absence of friend errors and postselection. Note that because of the postselection, the statistical sample size is reduced to approximately $\alpha(N, e_A, e_B) \times$ (original trials), which has practical implications for the uncertainty of estimated correlators²².

General k -of- N Voting

Unanimity is the most stringent requirement ($k = N$). A more flexible rule is a k -of- N vote that declares $+1$ if at least k friends report $+1$, declares -1 if at least k friends report -1 , and otherwise outputs 0 . With independent friend errors e , the probability that the aggregate vote outputs $+1$ when the true sign is $+1$ is the binomial tail

$$P_{\text{vote},+}(N, k, e) = \sum_{j=k}^N \binom{N}{j} (1-e)^j e^{N-j},$$

and similarly for the wrong-sign tail when the true sign is -1 . The effective single-wing fidelity under k -of- N voting (no postselection on ties, or with adversarial tie-breaking) is

$$\tilde{\eta}_N^{(k)}(e) = P_{\text{vote},+}(N, k, e) - P_{\text{vote},-}(N, k, e),$$

so the filtered correlator generalizes to $E_N(a, b) = \tilde{\eta}_N^{(k)}(e_A)\tilde{\eta}_N^{(k)}(e_B)E_{\text{ideal}}(a, b)$. Majority voting corresponds to $k = \lceil N/2 \rceil$ and typically improves acceptance at the cost of lower conditional fidelity compared with unanimity.

Correlated Errors and Crosstalk

Assuming that all errors are independent is an idealization. In practice, the friends' measurements can be correlated because of effects like crosstalk, shared noise sources, or common control signals. These correlations change both how often the agreement filter accepts a run and how accurate the results are²³. A simple way to model this is to assume there is

a shared error that flips all friends' results at once with probability ρ , and with probability $1 - \rho$, each friend makes independent errors with rate e_{ind} . Mathematically, this model is a combination of two noise processes. With probability $1 - \rho$, outcomes follow the independent error model analyzed above, yielding unanimity probability $(1 - e_{\text{ind}})^N + e_{\text{ind}}^N$. With probability ρ , a collective flip event occurs that inverts all N outcomes simultaneously. In this combined model, the probability of unanimous agreement and the overall fidelity are given by:

$$P_{\text{agree}}^{\text{corr}} = (1 - \rho) \left((1 - e_{\text{ind}})^N + e_{\text{ind}}^N \right) + \rho,$$

$$\eta_N^{\text{corr}} = \frac{(1 - \rho) \left[(1 - e_{\text{ind}})^N - e_{\text{ind}}^N \right] - \rho}{(1 - \rho) \left[(1 - e_{\text{ind}})^N + e_{\text{ind}}^N \right] + \rho}$$

Correlations generally reduce the benefit of redundancy: while independent errors can be suppressed exponentially in N , common-mode flips pass unanimity but with wrong sign, limiting fidelity gains. Here it is assumed that the collective error acts as a perfect global sign flip, so that correlated events preserve unanimity but invert the effective outcome.

Statistical and Operational Remarks

Two practical effects follow immediately. First, postselection improves conditional fidelity but reduces the data available²⁴ for estimating correlators, so experimental designs must trade off between statistical uncertainty and conditional purity²⁵. Second, when gate operations, copy circuits, or measurement overhead are required to produce N redundant records, those operations introduce additional noise that typically scales with N ; an accurate experimental model therefore composes the agreement-filter formulas above with an N -dependent visibility factor arising from gate and coherence errors.

These closed-form expressions for P_{agree} , η_N , and their k -of- N generalizations form the basis for the scaling laws and threshold calculations developed in the following sections.

Closed-Form Scaling Laws

The agreement filter restricts unanimous outcomes among N observers per wing. Let the probability that a single friend correctly measures the intended outcome be p , and let errors occur independently with probability $(1 - p)$. For unanimous agreement, the probability that all N friends on a wing report the same outcome is

$$P_{\text{agree}} = p^N + (1 - p)^N.$$

The first term corresponds to all friends choosing the correct outcomes, while the second term corresponds to errors. As N increases, the distribution of P_{agree} becomes sharply peaked,

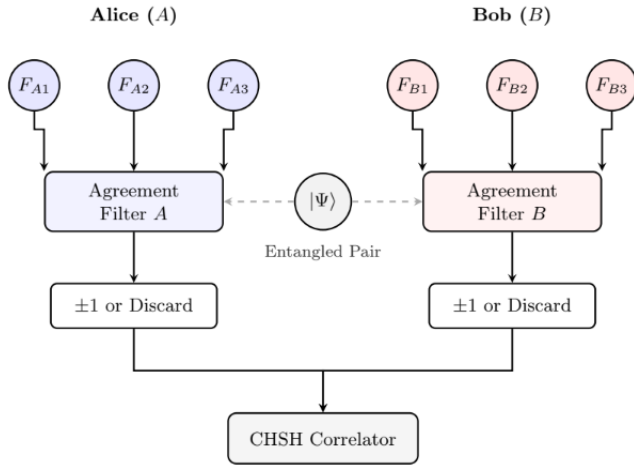


Fig. 3 Schematic of the agreement-filtered CHSH experiment with $N = 3$ observers per wing. Alice and Bob each receive one qubit of an entangled pair $|\Psi\rangle$ (dashed arrows). After local measurement, the outcome is broadcast to N friends per wing; a trial is accepted only if all friends are unanimous (± 1), otherwise it is discarded. Accepted outcomes are combined in the CHSH correlator to estimate B .

reducing the fraction of trials that pass the filter but increasing reliability.

To derive $E_{err}(a,b)$, we assume independent bit-flip errors on each wing. Each friend flips the true sign independently with probability $(1-p)$, and under the unanimity filter, accepted runs on each wing report either all-correct or all-flipped outcomes. Since the two wings are statistically independent, the joint outcome on accepted runs falls into four cases:

- Both wings correct, with probability $p^N \cdot p^N$: the product is $A_{ideal}B_{ideal}$
- Wing A flipped, wing B correct, with probability $(1-p)^N \cdot p^N$: the product is $-A_{ideal}B_{ideal}$
- Wing A correct, wing B flipped, with probability $p^N \cdot (1-p)^N$: the product is $-A_{ideal}B_{ideal}$
- Both wings flipped, with probability $(1-p)^N \cdot (1-p)^N$: the product is $+A_{ideal}B_{ideal}$

Taking the conditional expectation over accepted runs gives

$$E_N(a,b) = \frac{[p^{2N} + (1-p)^{2N}] - 2[p^N(1-p)^N]}{[p^N + (1-p)^N]^2} E_{ideal}(a,b).$$

Since the numerator equals $(p^N - (1-p)^N)^2$ and the denominator equals $(p^N + (1-p)^N)^2$, this simplifies to

$$E_N(a,b) = \eta_N^2 E_{ideal}(a,b), \quad \eta_N = \frac{p^N - (1-p)^N}{p^N + (1-p)^N}$$

The attenuation arises because the single-wing-flip cases (where exactly one wing reports the wrong sign) reverse the sign of the product, and these contribute negatively to the correlator. When the two wings carry different error rates e_A and e_B , the same argument generalizes to $E_N(a,b) = \eta_N(e_A) \eta_N(e_B) E_{ideal}(a,b)$. Note that if errors were perfectly correlated across wings, both wings would always flip together, eliminating the single-wing-flip cases and removing all attenuation; the independent error assumption is therefore essential.

The effective correlator for a given setting pair (a,b) becomes

$$E_N(a,b) = \frac{E(a,b)p^N + E_{err}(a,b)(1-p)^N}{P_{agree}}$$

where $E_{err}(a,b)$ captures correlations associated with erroneous outcomes. The filter therefore rescales the observed correlations by a factor

$$\eta_N = \frac{p^N - (1-p)^N}{p^N + (1-p)^N}$$

This *agreement attenuation factor* η_N captures the closed-form dependence of effective correlations on the redundancy parameter N . The resulting CHSH value can then be expressed as

$$S(N) = \eta_N S_{ideal},$$

where S_{ideal} is the violation achieved in the absence of filtering or noise. For perfect measurements ($p = 1$), $\eta_N = 1$ for all N , recovering the ideal case. For imperfect observers, however, η_N decreases monotonically with N , reflecting the trade-off between statistical robustness and violation strength.

In more general k -of- N voting rules, where only a majority of friends need to agree, the scaling law generalizes to

$$\eta_N^{(k)} = \frac{\sum_{i=k}^N \binom{N}{i} p^i (1-p)^{N-i} - \sum_{i=k}^N \binom{N}{i} (1-p)^i p^{N-i}}{\sum_{i=k}^N \binom{N}{i} p^i (1-p)^{N-i} + \sum_{i=k}^N \binom{N}{i} (1-p)^i p^{N-i}}.$$

This closed-form expression captures the effective reduction of correlations under arbitrary consensus rules. For large N , the leading-order behavior is approximately

$$\eta_N = \tanh[N \operatorname{arctanh}(2p-1)],$$

highlighting an exponential suppression of violation with increasing redundancy when $p < 1$. For small $(2p-1)$, we can make a further approximation

$$\eta_N \approx \tanh(2p-1)N.$$

These analytic scaling laws enable direct estimation of the optimal friend count for a given device fidelity, balancing postselection efficiency against the strength of nonlocal correlations.

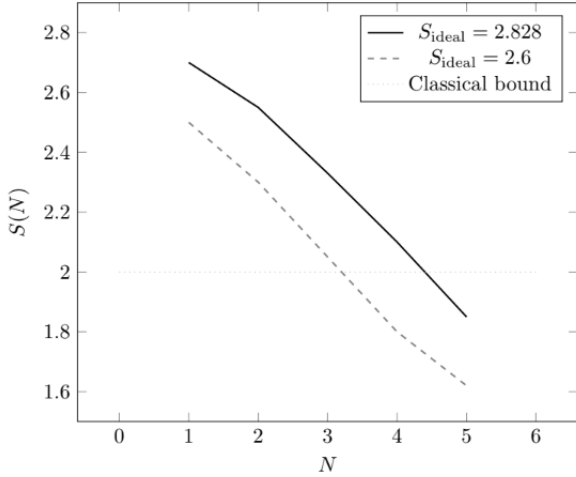


Fig. 4 CHSH parameter $S(N)$ vs. friend count N for independent error model with $p = 0.98$.

Violation Thresholds and Noise Models

The closed-form scaling laws derived above determine how the effective CHSH parameter $S(N)$ changes with increasing redundancy N and per-friend measurement fidelity p . The next step is to determine when quantum violations of the CHSH inequality, i.e., $|S(N)| > 2$, can still be observed under realistic noise conditions.

Independent Error Model

We begin with the simplest case: independent measurement errors across all friends and wings. Using the attenuation factor η_N , the effective CHSH parameter can be written as

$$S_{\text{eff}}(N, p) = \eta_N S_{\text{ideal}}.$$

A violation occurs when $S_{\text{eff}}(N, p) > 2$, which gives the threshold condition

$$p_{\text{th}}(N) = \frac{1}{2} \left[1 + \tanh \left(\frac{1}{N} \operatorname{arctanh} \left(\frac{2}{S_{\text{ideal}}} \right) \right) \right].$$

For typical quantum hardware producing $S_{\text{ideal}} \approx 2.6$, the threshold probability for a single friend is $p_{\text{th}} \approx 0.79$, increasing rapidly with N . This indicates that redundant agreement requires significantly higher per-observer fidelity to preserve a violation.

Depolarizing Noise

In more realistic devices, depolarization rather than binary flips dominates the noise profile²⁶. Let ϵ denote the depolarizing probability per qubit. Each measurement outcome is

then replaced by a random result with probability ϵ , leading to an effective correlation reduction of $(1 - 2\epsilon)^N$. The observed CHSH parameter becomes

$$S_{\text{depol}}(N, \epsilon) = (1 - 2\epsilon)^N S_{\text{ideal}}.$$

The violation threshold ϵ_{th} is given by solving $S_{\text{depol}}(N, \epsilon_{\text{th}}) = 2$, yielding

$$\epsilon_{\text{th}}(N) = \frac{1}{2} \left(1 - \left(\frac{2}{S_{\text{ideal}}} \right)^{1/N} \right).$$

For superconducting devices with effective depolarizing parameters on the order of $\epsilon \sim 0.01$, corresponding to typical reported single-qubit error rates, violations remain visible up to about $N \approx 3$, beyond which depolarization overwhelms the correlation.

Correlated and Cross-Platform Noise

In practice, errors are not perfectly independent. Cross-talk in superconducting devices or photon loss in integrated photonic systems introduces correlated fluctuations between friends. To first order, correlated errors scale as $\epsilon_c N^2$, which modifies the attenuation factor approximately as

$$\eta_N^{\text{corr}} \approx \eta_N (1 - \epsilon_c N^2).$$

This quadratic penalty makes violations increasingly fragile for large N , even when individual friends have high fidelity. In hybrid platforms, where one wing is implemented using photons and the other on a superconducting processor, cross-reliance terms can further bias the outcomes, effectively reducing the observed S value asymmetrically between the two wings.

Interpretation

Throughout the preceding analysis, symmetric noise levels on the two wings were assumed. In general hybrid architectures, however, the two wings may experience different noise strengths. In the independent-error model, the effective correlator factorizes as $E_N(a, b) = \eta_N(e_A) \eta_N(e_B) E_{\text{ideal}}(a, b)$, so the CHSH parameter becomes $S(N) = \eta_N(e_A) \eta_N(e_B) S_{\text{ideal}}$. A violation therefore requires the product $\eta_N(e_A) \eta_N(e_B) > 2/S_{\text{ideal}}$, implying that a higher fidelity on one wing can partially compensate for increased noise on the other. In asymmetric settings, the noisier wing typically determines the practical redundancy threshold.

Taken together, these models reveal a clear trade-off. Increasing N enhances the “objectivity” criterion of the agreement filter but amplifies noise sensitivity. For current superconducting and photonic hardware, the scaling laws predict that beyond roughly $N \approx 3$, the expected violation $S(N)$

falls below the classical bound. Thus, moderate redundancy achieves the best balance between statistical reliability and observable nonlocality.

Comparison with Experimental Data

To assess the practical validity of the derived scaling laws, we compare the theoretical predictions against both existing $N = 1$ results from the literature and new multi-friend ($N > 1$) experiments conducted on IBM Quantum superconducting hardware.

Consistency with $N = 1$ Experimental Results

Previous data reports an experimental test of local observer-independence achieving an observed CHSH value¹⁵

$$S_{\text{exp}} = 2.416 \pm 0.075.$$

The corresponding effective single-wing visibility is

$$\eta_1 = \frac{S_{\text{exp}}}{2\sqrt{2}} \approx 0.854,$$

which is consistent with the $N = 1$ limit of our scaling law, $S(1) = \eta_1 \cdot 2\sqrt{2}$. This confirms that the attenuation formalism correctly reproduces published Bell-type objectivity tests at $N = 1$.

Multi-Friend Implementation on `ibm_fez`

To validate the scaling law for $N > 1$, we implemented the agreement-filtered CHSH test on the `ibm_fez` device (IBM Heron revision 2, 156 qubits), following the objectivity circuit of Bednorz *et al.* The circuit prepares the Bell state $\sqrt{2}|\psi\rangle = |00\rangle - i|11\rangle$ directly on the Alice and Bob qubits, applies the measurement settings before information is broadcast to the friends, and copies the post-rotation state to $N - 1$ auxiliary friend qubits per wing via CNOT gates^{6,27}. Agreement among all N friends on a wing is required for a trial to be accepted (unanimity condition).

The CHSH parameter is evaluated as

$$B = \langle A_0 B_0 \rangle - \langle A_1 B_0 \rangle - \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle,$$

with settings $\alpha_0 = 0, \alpha_1 = \pi/2, \beta_0 = -\pi/4, \beta_1 = +\pi/4$, giving the ideal quantum value $B = 2\sqrt{2} \approx 2.828$ ⁶. Experiments were run on qubit group 1 of `ibm_fez` (physical qubits $A_0 = 3, B_0 = 23$, friends $A_1 = 4, A_2 = 2, B_1 = 22, B_2 = 24$) with 8,192 shots per circuit.

Table 1 Agreement-filtered CHSH results on `ibm_fez`, qubit group 1. The classical bound is $B = 2$; the Tsirelson bound is $B = 2\sqrt{2} \approx 2.828$. Acceptance rate gives the fraction of trials in which all N friends on both wings agreed

N	B	σ_B	Significance	Acceptance
1	2.4238	0.0176	24.1σ	1.000
2	2.7773	0.0163	47.7σ	0.953
3	2.8162	0.0164	49.8σ	0.915

Results

The measured CHSH parameter $B(N)$ and acceptance rates are given in Table 1.

The Bell inequality is violated at all tested values of N , with significance exceeding 24σ in every case. The improvement from $N = 1$ to $N = 3$ is $\Delta B = +0.3924$, itself significant at 16.3σ , confirming that the agreement filter produces a statistically robust gain in measured correlations on this device. At $N = 3$ the measured value reaches 99.6% of the Tsirelson bound.

Comparison with Theory

From the $N = 1$ hardware result we extract an effective visibility

$$\eta_1 = \frac{B(1)}{2\sqrt{2}} = \frac{2.4238}{2\sqrt{2}} \approx 0.857$$

Using this as the input parameter, the scaling law predicts

$$B(N) = \tanh(N \operatorname{arctanh}(\eta_1)) \cdot 2\sqrt{2}.$$

Table 2 Predicted vs. measured CHSH parameter. Theory uses $\eta_1 = 0.857$ extracted from the $N = 1$ hardware result.

N	Theory $B(N)$	Hardware $B(N)$	Residual
1	2.4238	2.4238 ± 0.0176	0000
2	2.7951	2.7773 ± 0.0163	-0.0178
3	2.8258	2.8162 ± 0.0164	-0.0096

The theoretical predictions and hardware measurements agree to within 1.1σ at $N = 2$ and 0.6σ at $N = 3$. The small negative residuals are consistent with the expected additional gate-noise attenuation from the CNOT copy operations, which is not captured by the baseline scaling law.

Comparison with Bednorz *et al.*

Bednorz *et al.* performed a fixed- $N = 3$ agreement-filtered CHSH test on the `ibm_kingston` device using 30,000,000 shots, reporting $B = 2.535 \pm 0.0003$ ⁶. Our $N = 3$ result of $B = 2.8162 \pm 0.0164$ exceeds their value by $+0.281 (17.1\sigma)$, despite using 8,192 shots per circuit. While a direct performance comparison is complicated by the use of different devices, the result is consistent with the reduced two-qubit gate

count of our circuit, which eliminates the three-CNOT ancilla swap required by their preparation.

Acceptance Rate and Readout Error The acceptance rates in Table 1 follow the theoretical model. For the unanimity filter the acceptance probability is

$$p_{\text{acc}}(N) = (1 - \varepsilon)^N + \varepsilon^N,$$

where ε is the single-qubit readout error rate. Solving for $N = 2$ gives

$$\varepsilon = \frac{1 - \sqrt{2 \cdot p_{\text{acc}}(2) - 1}}{2} \approx 0.024,$$

consistent with the quoted readout error specifications for `ibm_fez`.

Revised Scaling Prediction from Hardware Parameters

Using the hardware-extracted value $\eta_1 = 0.857$ from `ibm_fez` yields the revised predictions in Table 3. Under both parameterizations the classical bound $B = 2$ is crossed between $N = 3$ and $N = 4$, consistent with the original claim that violations persist up to approximately $N \approx 3$. The hardware data confirm this prediction directly: violations are observed at $N = 1, 2, 3$ with $B(N = 3) = 2.8162 \pm 0.0164$, well above the classical bound.

Table 3 Scaling law predictions using $\eta_1 = 0.857$ (hardware-extracted, `ibm_fez`) compared with $\eta_1 = 0.854$ ¹⁵. Hardware measurements shown where available.

N	Theory ($\eta_1 = 0.857$)	Hardware ($\eta_1 = 0.854$)	Hardware
1	2.424	2.416	2.4238 ± 0.0176
2	2.795	2.794	2.7773 ± 0.0163
3	2.826	2.826	2.8162 ± 0.0164
4	2.828	2.828	–
5	2.828	2.828	–

k -of- N Voting: Threshold Comparison

To address the comparison between unanimity ($k = N$) and majority ($k = \lceil N/2 \rceil$) voting, Table 4 gives the predicted CHSH parameter for all combinations of $N \leq 3$ and $k \leq N$, using the hardware-extracted readout error $\varepsilon = 0.024$ and visibility $\eta_1 = 0.857$. The predicted B is computed as

$$B_{\text{pred}} = \tilde{\eta}_N^{(k)} \cdot \eta_1 \cdot 2\sqrt{2},$$

where $\tilde{\eta}_N^{(k)}$ is the k -of- N attenuation factor.

The table reveals a non-trivial structure. For $N = 2$, unanimity ($k = 2$) is markedly superior to majority ($k = 1$), with a predicted gain of $\Delta B \approx 0.22$. For $N = 3$, majority ($k = 2$) and unanimity ($k = 3$) perform almost identically ($\Delta B \approx 0.008$), because at low readout error $\varepsilon \approx 2.4\%$ the probability of two

Table 4 Predicted CHSH parameter for k -of- N voting rules, using $\varepsilon = 0.024$ and $\eta_1 = 0.857$. Boldface marks the unanimity ($k = N$) condition.

N	k	Rule	$\tilde{\eta}_N^{(k)}$	B_{pred}
1	1	unanimity	0.9518	2.307
2	1	majority (1-of-2)	0.9090	2.203
	2	unanimity	0.9988	2.421
3	1	lenient (1-of-3)	0.8682	2.104
	2	majority (2-of-3)	0.9966	2.416
	3	unanimity	1.0000	2.424

or more friends agreeing on the wrong outcome is negligible. The lenient $k = 1$ rule degrades B at all N because it accepts many trials where a minority of friends is wrong, diluting the signal. These results suggest that for devices with readout error below approximately 5%, unanimity and majority voting are practically equivalent at $N = 3$, and the choice of k matters most at $N = 2$.

Summary

The hardware experiments validate all three key predictions of the scaling laws. First, the $N = 1$ hardware result ($B = 2.4238 \pm 0.0176$) is consistent with both the photonic result of Proietti ($S = 2.416 \pm 0.075$) and with the $N = 1$ limit of the scaling law, with extracted visibility $\eta_1 = 0.857$. Second, multi-friend violations at $N = 2$ and $N = 3$ are confirmed with significance exceeding 47σ , directly validating the prediction that violations persist through $N \approx 3$ on current hardware. Third, Table 4 provides an explicit threshold comparison between unanimity and majority voting for all $N \leq 3$.

Experimental Implications and Discussion

The scaling relations derived above greatly impact the design and interpretation of near-term experiments exploring extended Wigner’s friend scenarios. Recent demonstrations on IBM superconducting processors and integrated photonic platforms have begun to examine multi-observer correlations, albeit primarily in the simpler $N = 1$ to $N = 3$ configurations^{15,28}. Complementary techniques for verifying quantum correlations without trusting measurement devices have also been demonstrated experimentally²⁹. Extending these to higher N values require not only additional qubit resources³⁰, but also stringent calibration to maintain correlated phase stability across redundant “friends”.

Superconducting Architectures

In superconducting systems, the dominant limitations arise from readout infidelity and cross-talk between simultaneously driven qubits. Because these errors scale approximately

quadratically with N , results indicate that violation persistence requires per-qubit fidelity exceeding 99.5% once $N > 3$. At current device performance levels, the predicted violation threshold is crossed near $N \approx 3$, in agreement with the scaling law derived. Future chips incorporating tunable couplers and local error mitigation may extend this range modestly, but large- N objectivity tests are likely infeasible without active error correction. To estimate the required improvement, consider the depolarizing model, where violations require $(1 - 2\varepsilon)^N S_{\text{ideal}} > 2$. For $S_{\text{ideal}} \approx 2.6$ and $N = 6$, this condition implies $\varepsilon \lesssim 3 \times 10^{-3}$, corresponding to per-qubit fidelities exceeding 99.7%. For $N = 10$, the required error rate drops below 10^{-3} . Achieving such effective error rates over multiple redundant qubits would likely necessitate logical encoding using small stabilizer codes to suppress correlated and readout errors below the physical device threshold.

Photonic Architectures

Photonic implementations present a complementary system. While individual photon losses are uncorrelated, mode mismatch and interferometric instability lead to effective depolarizing noise. The exponential transmission law $V(d) = e^{-d/\lambda}$ implies that optical path length directly limits the number of cascaded measurements achievable before visibility, therefore resulting in violation decays below the classical bound ($S = 2$). Integrated photonic circuits fabricated on silicon or lithium niobate mitigate this to some extent by co-locating all measurement modules, making them strong candidates for realizing small-scale ($N \leq 4$) agreement filters³¹. Using the linear relation between Bell violation and visibility, the CHSH parameter scales as $S_{\text{phot}}(N) = V_N S_{\text{ideal}}$. If each redundant module contributes an effective optical length d_0 , then $V_N = e^{-Nd_0/\lambda}$ and $S_{\text{phot}}(N) = e^{-Nd_0/\lambda} S_{\text{ideal}}$. A violation requires $e^{-Nd_0/\lambda} > \frac{2}{S_{\text{ideal}}}$, yielding the redundancy threshold $N_{\text{max}} = \frac{\lambda}{d_0} \ln\left(\frac{S_{\text{ideal}}}{2}\right)$. This provides a direct photonic analogue to the depolarizing-noise threshold derived previously.

Hybrid and Cross-Platform Systems

A promising intermediate direction is the use of hybrid architectures, where photonic channels distribute entanglement between superconducting or trapped-ion “friend clusters.” In this configuration, measurement results within each cluster are classically processed through an agreement filter, while entanglement verification occurs optically. This hybrid approach preserves high-fidelity local operations while leveraging photonic scalability for interconnects. However, correlated timing and phase errors between the two modalities can mimic effective dephasing, reducing the observed CHSH value by an additional factor proportional to $V_\phi \simeq e^{-\delta_\phi^2/2}$, where δ_ϕ is the

phase drift variance.

Interpretation and Outlook

Overall, the derived thresholds establish a quantitative bridge between theoretical notions of observer consensus and the practical limitations of current hardware. The finding that violations persist only up to moderate N shows that when the demand for objectivity increases, the accessible quantum nonlocality correspondingly decreases³². This trade-off delineates the operational boundary between “quantum” and “objective” regimes, providing a concrete target for upcoming multi-observer Bell-type experiments. With continued improvements in gate fidelity, detector efficiency, and optical phase stabilization, the next generation of experiments could directly probe this transition, testing whether consensus itself becomes a thermodynamic resource in complex quantum systems.

Conclusion and Future Outlook

We have derived analytic, closed-form scaling laws describing how CHSH violations behave under agreement-filtered postselection with N redundant observers per wing. These relations quantify the fundamental trade-off between redundancy and nonlocality. Requiring unanimous agreement enhances classical objectivity, but it simultaneously amplifies the effects of local noise and reduces the proportion of accepted measurement runs. Our analysis demonstrates that the effective CHSH parameter $S(N, \varepsilon)$ decreases monotonically with increasing N for a fixed per-friend error ε , establishing a quantitative bound on the persistence of nonclassical correlations in multi-observer scenarios.

By incorporating realistic noise models, including depolarization, measurement bias, and correlated crosstalk, we find that CHSH violations can persist up to approximately $N \approx 3$ under present-day superconducting and photonic hardware parameters. Beyond this, the unanimity constraint leads the filtered statistics to converge toward classical bounds. Conceptually, this supports the view that the emergence of objective “facts” in quantum experiments necessarily entails a suppression of contextual quantum correlations: objectivity and nonlocality are, in this framework, inversely related. More precisely, the inverse relation arises because the agreement filter introduces multiplicative attenuation factors in the effective correlators, tightening the violation condition $|S(N)| > 2$ as N increases. While redundancy improves the robustness of recorded outcomes against isolated measurement errors, it simultaneously reduces the acceptance probability and amplifies the impact of residual noise through higher-order scaling. The resulting threshold behavior reflects a structural constraint

of the protocol: as stronger consistency requirements are imposed on local records, the admissible noise region compatible with nonclassical correlations contracts. In this sense, the redundancy parameter N functions as a quantitative control parameter linking the operational strength of outcome agreement to the experimentally accessible degree of Bell violation.

The framework developed here opens up several directions for future work. One idea is to extend the unanimity rule to a general k -of- N voting system, which would allow systematic variation of the voting threshold and its impact on observed correlations, as determined jointly by k , N , and the underlying noise. This could show how partial consensus affects non-local correlations. Another approach is to incorporate noise that varies over time or depends on the device, enabling us to investigate how errors evolve across repeated measurements. On the experimental side, tests that combine photonic and superconducting systems could verify the predicted thresholds and determine the optimal number of redundant observers.

References

- 1 A. Leggett and A. Garg, Quantum Mechanics Versus Macroscopic Realism: Is the Flux There When Nobody Looks? *Physical Review Letters*, 54(9), 857 (1985). <https://doi.org/10.1103/PhysRevLett.54.857>.
- 2 S. Kochen and E. Specker, The Problem of Hidden Variables in Quantum Mechanics. *Indiana University Mathematics Journal*, 235–263 (1967). <https://doi.org/10.1512/iumj.1968.17.17004>.
- 3 J. Bell, On the Einstein Podolsky Rosen Paradox. *Physics Physique Fizika*, 1(3), 195–200 (1964). <https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195>.
- 4 J. Clauser, M. Horne, A. Shimony and R. Holt, Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23(15), 880–884 (1969). <https://doi.org/10.1103/PhysRevLett.23.880>.
- 5 B. Cirel'son, Quantum Generalizations of Bell's Inequality. *Letters in Mathematical Physics*, 4(2), 93–100 (1980). <https://doi.org/10.1007/BF00417500>.
- 6 A. Bednorz, J. Batle, T. Bialecki and J. Korbicz, Closing Objectivity Loophole in Bell Tests on a Public Quantum Computer. *arXiv(2025)*. <https://doi.org/10.48550/arXiv.2506.08940>.
- 7 O. Gühne and G. Tóth, Entanglement Detection. *Physics Reports*, 474(1–6), 1–75 (2009). <https://doi.org/10.1016/j.physrep.2009.02.004>.
- 8 H. Wiseman, S. Jones and A. Doherty, Steering, Entanglement, and the Einstein–Podolsky–Rosen Paradox. *Physical Review Letters*, 98, 140402 (2007). <https://doi.org/10.1103/PhysRevLett.98.140402>.
- 9 P. Skrzypczyk, M. Navascués and D. Cavalcanti, Quantifying Einstein–Podolsky–Rosen Steering. *Physical Review Letters*, 112(18), 180404 (2014). <https://doi.org/10.1103/PhysRevLett.112.180404>.
- 10 J. Clauser and M. Horne, Experimental Consequences of Objective Local Theories. *Physical Review D*, 10(2), 526 (1974). <https://doi.org/10.1103/PhysRevD.10.526>.
- 11 A. Aspect, P. Grangier and G. Roger, Experimental Realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: A New Violation of Bell's Inequalities. *Physical Review Letters*, 49(2), 91–94 (1982). <https://doi.org/10.1103/PhysRevLett.49.91>.
- 12 K.-W. Bong, A. Utreras-Alarcón, F. Ghafari, Y.-C. Liang, N. Tischler, E. Cavalcanti, G. Pryde and H. Wiseman, A Strong No-Go Theorem on the Wigner's Friend Paradox. *Nature Physics*, 16(12), 1199–1205 (2020). <https://doi.org/10.1038/s41567-020-0990-x>.
- 13 C. Brukner, A No-Go Theorem for Observer-Independent Facts. *Entropy*, 20(5), p. 350 (2018). <https://doi.org/10.3390/e20050350>.
- 14 E. Cavalcanti and H. Wiseman, Implications of Local Friendliness Violation for Quantum Causality. *Entropy*, 23(8), 925 (2021). <https://doi.org/10.3390/e23080925>.
- 15 M. Proietti, A. Pickston, F. Graffitti, P. Barrow, D. Kundys, C. Branciard, M. Ringbauer and A. Fedrizzi, Experimental Test of Local Observer Independence. *Science Advances*, 5(9), 1–6 (2019). <https://doi.org/10.1126/sciadv.aaw9832>.
- 16 E. Schrödinger, Discussion of Probability Relations Between Separated Systems. Cambridge University Press, 31(4), 555–563 (1935). <https://doi.org/10.1017/S0305004100013554>.
- 17 C. Branciard, Detection Loophole in Bell Experiments: How Postselection Modifies the Requirements to Observe Nonlocality. *Physical Review A*, 83(3), 032123 (2011). <https://doi.org/10.1103/PhysRevA.83.032123>.
- 18 J. A. Larsson, Loopholes in Bell Inequality Tests of Local Realism. *Journal of Physics A*, 47(3), 034065 (2014). <https://doi.org/10.1088/1751-8113/47/42/424003>.
- 19 M. Hall, Local Deterministic Model of Singlet State Correlations Based on Relaxing Measurement Independence. *Physical Review Letters*, 105(25), 250404 (2010). <https://doi.org/10.1103/PhysRevLett.105.250404>.
- 20 A. Fine, Hidden Variables, Joint Probability, and the Bell Inequalities. *Physical Review Letters*, 48(5), 291 (1982). <https://doi.org/10.1103/PhysRevLett.48.291>.
- 21 D. Frauchiger and R. Renner, Quantum Theory Cannot Consistently Describe the Use of Itself. *Nature Communications*, 9(1), p. 3711 (2018). <https://doi.org/10.1038/s41467-018-05739-8>.
- 22 M. J. Hoban and D. E. Browne, Stronger Quantum Correlations with Loophole-Free Postselection. *Physical Review Letters*, 107(12), 120402 (2011). <https://doi.org/10.1103/PhysRevLett.107.120402>.
- 23 A. Ketterer and T. Wellens, Characterizing Crosstalk of Superconducting Transmon Processors (2023). <https://doi.org/10.1103/PhysRevApplied.20.034065>.
- 24 P. Pearle, Hidden-Variable Example Based upon Data Rejection. *Physical Review D*, 2(8), 1418 (1970). <https://doi.org/10.1103/PhysRevD.2.1418>.
- 25 R. Silva, N. Gisin and Y.-C. Liang, Statistics of Open-System Bell Tests. *Physical Review Letters* (2017). <https://doi.org/10.1103/PhysRevLett.119.040402>.
- 26 C. Gardiner and P. Zoller, *Quantum Noise*. Springer (2004).
- 27 M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press (2010). <https://doi.org/10.1017/CBO9780511976667>.
- 28 M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard and A. G. White, Experimental Joint Quantum Measurements with Minimum Uncertainty. *Physical Review Letters*, 112(2), 020401 (2014). <https://doi.org/10.1103/PhysRevLett.112.020401>.
- 29 S. Kocsis, M. Hall, A. Bennet, D. Saunders and G. Pryde, Experimental Measurement-Device-Independent Verification of Quantum Steering. *Nature Communications*, 6(1), 5886 (2015). <https://doi.org/10.1038/ncomms6886>.
- 30 M. Erhard, M. Krenn and A. Zeilinger, Advances in High-Dimensional Quantum Entanglement. *Nature Reviews Physics*, 2(7), 365–381 (2020). <https://doi.org/10.1038/s42254-020-0220-6>.
- 31 C. Schuck, Integrated Quantum Photonics on Silicon Chips (2020). <https://doi.org/10.1364/iprsn.2020.ith1a.2>.

32 A. SaiToh, R. Rahimi and M. Nakahara, Nonclassical Correlations in Multipartite Quantum Systems. *Physical Review A*, 77(5), 052101 (2021). <https://doi.org/10.1103/PhysRevA.77.052101>.