

Entanglement certification by measuring nonlocality

Xuan Du Trinh

Department of Computer Science
Stony Brook University, USA
xtrinh@cs.stonybrook.edu

Zhengyu Wu

Department of Computer Science
Stony Brook University, USA
zhenwu@cs.stonybrook.edu

Junlin Bai

Department of Computer Science
Stony Brook University, USA
jubbai@cs.stonybrook.edu

Huan-Hsin Tseng

AI & ML Department
Brookhaven National Laboratory, NY, USA
htseng@bnl.gov

Nengkun Yu

Department of Computer Science
Stony Brook University, USA
nengkun.yu@cs.stonybrook.edu

Aruna Balasubramanian

Department of Computer Science
Stony Brook University, USA
arunab@cs.stonybrook.edu

Abstract—Reliable entanglement manipulation is key to quantum networks. This paper presents a practical approach to entanglement verification based on Clauser–Horne–Shimony–Holt (CHSH) inequality violation [1]. We establish two-sided bounds relating the CHSH statistic to the target-state fidelity, and develop a statistical framework for estimating $S(\rho)$ from finite measurements with explicit sample complexity and confidence guarantees. Our main contributions include a CHSH-based Entanglement Verification (EV) protocol with worst-case type-I and type-II error control, Monte Carlo simulations on depolarized Werner states showing that the prescribed sample sizes are substantially conservative in practice, and a two-node network simulation in SeQUeNCe [2] showing how channel degradation and memory decoherence shift the effective verification-time fidelity below the source fidelity. These results illustrate the feasibility of CHSH-based verification as a link-level primitive in quantum networks and identify the source-fidelity margin needed to maintain downstream teleportation quality.

Index Terms—Quantum Networking, Entanglement Verification, Nonlocality

I. INTRODUCTION

Quantum networks are emerging as a key infrastructure for secure communication and distributed quantum information processing, enabling applications such as quantum key distribution [3], [4], delegated quantum computation [5], [6], and long-distance quantum communication [7]–[9]. The standard mechanism for transmitting quantum information over a network is *quantum teleportation*, which uses a shared Einstein–Podolsky–Rosen (EPR) pair $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ together with classical communication to transfer an arbitrary quantum state from a sender to a receiver [7].

In practice, generating high-fidelity EPR pairs over long distances is difficult. Decoherence, photon loss, and imperfect operations degrade entanglement [10]–[12]. Teleportation over low quality entanglement results in loss of the transmitted qubit. The problem is exacerbated by the no-cloning theorem which prevents the quantum state from being copied or retransmitted. If the qubit is lost during teleportation, it will have to be regenerated.

Several approaches have been proposed to improve the quality of entanglement, including error-correction-based repeater architectures [13], and entanglement purification, which uses local operations and classical communication to distill higher-fidelity pairs from noisier ones [14], [15]. These techniques improve the *average* quality of distributed entanglement.

However, the fundamental problem remains: while purification and error-correction protocols can improve entanglement fidelity, they cannot guarantee that the resulting pairs are of sufficiently high quality. This motivates the need for *entanglement verification*: verifying that the entangled pair has high fidelity *before* using it for teleportation. If we can verify that an EPR pair has high enough fidelity for teleportation, they are used for communication; otherwise the pair should be discarded and a different pair generated.

The problem is that existing entanglement verification techniques have been studied extensively as a standalone procedure, but are not well-suited in the context of quantum communication. Several existing entanglement verification algorithms are destructive [16]–[19], in that, even though the algorithm can verify the fidelity, the entanglement collapses at the end of the verification process. This means that the verified entanglements can no longer be used for teleportation. Verification using entanglement witnesses [20]–[23] and control unitaries [24] focus on binary verification; they can either guarantee that the qubits are maximally entangled or are orthogonal. However, in quantum networks, qubits are rarely maximally entangled, but teleportation is still possible over these imperfect entanglements. Other techniques such as quantum state tomography provides complete state reconstruction but incurs prohibitive resource overhead [25]–[27].

In this work, we propose an efficient and non-destructive entanglement verification approach that can be integrated with quantum communication. The core idea is to leverage violation of the CHSH inequality [1]. The violation certifies the presence of entanglement, and the degree of violation should directly reflect the level of entanglement: the stronger the violation, the higher the fidelity with the ideal EPR pair.

Moreover, CHSH is a natural fit for networks that already perform Bell tests: the detector calibration, timing, and

classical post-processing infrastructure required for CHSH measurements are already in place in networks running device-independent quantum key distribution (DIQKD) [28], [29]. Under trusted and fixed local measurement settings, we show that the CHSH value yields explicit analytical lower and upper bounds on the target-state fidelity $F(\rho)$: a sufficiently large value certifies that the EPR pair has high fidelity, while a small value certifies that it is inadequate for teleportation.

However, CHSH measurement is inherently destructive: measuring the pair to estimate the CHSH value collapses the quantum state, consuming the pair and making it unavailable for teleportation. To address this, we adopt a *batch verification framework* in which each pair of communication nodes establishes a *batch* of EPR pairs. The parties sacrifice a fraction of their distributed pairs [30] to estimate the CHSH value, and construct a confidence interval on this estimate. The resulting bounds on $F(\rho)$ then drive a statistically controlled go/no-go decision before the remaining pairs are committed to teleportation.

Our main contributions are as follows.

- 1) **CHSH-based entanglement verification.** We adopt a batch-level entanglement verification framework in which a fraction of distributed pairs is sacrificed for CHSH measurements, and the result is used to make a statistically controlled go/no-go decision. To this end, we derive closed-form lower and upper bounds relating the CHSH value to the target-state fidelity $F(\rho)$ under fixed trusted measurements, providing the analytical foundation for the go/no-go decision.
- 2) **Statistical framework and verification protocol.** We quantify how many EPR pairs need to be sacrificed to estimate the CHSH value within a prescribed accuracy and confidence. Building on this estimate, our entanglement verification protocol distinguishes between a high-fidelity regime and a low-fidelity regime, with explicit worst-case type-I and type-II error bounds and sample complexity that scales with the desired precision and confidence.
- 3) **Evaluations using network simulation and analysis.** Using SeQUeNCe [2], we build a two-node simulation framework and implement our entanglement verification protocol. This simulation platform allows us to take into account realistic channel loss, delay, and memory decoherence. We show the effect of varying noise, distance, memory decoherence and the number of measurements on the performance of the protocol. Our experiments show that our approach is able to reliably verify if a batch of EPR pairs have higher than a threshold target fidelity.

II. RELATED WORKS

Research on entanglement verification spans several complementary directions. We review these works and discuss why they are insufficient in the context of verification for quantum teleportation.

Quantum state tomography and direct fidelity estimation. Quantum state tomography (QST) reconstructs the full density matrix and therefore gives complete information about the underlying state [25]–[27]. Its main drawback is measurement cost, which becomes prohibitive as systems grow. Direct fidelity estimation (DFE) provides a more scalable alternative when the target state is known [31], [32]. Rather than reconstructing the full density matrix, DFE estimates fidelity by sampling a tailored subset of Pauli observables, and its sample cost depends on the desired statistical accuracy rather than the system dimension. DFE is a strong baseline for target-state fidelity estimation. However, DFE is destructive; after verification, one can no longer use the EPR for teleportation and is therefore not well-suited in the context of quantum communication. We note that the CHSH method we propose in this work, on its own, is also destructive. However, we design a statistical verification technique that sacrifices a fraction of EPR pairs to non-destructively verify the remaining pairs. In addition, DI-QKD has already been experimentally demonstrated [28], [29], meaning that networks running it already have the detector calibration, timing, and classical post-processing infrastructure required for CHSH measurements, making CHSH a natural fit for practical deployment.

Entanglement witnesses and recent verification methods. Entanglement witnesses (EWs) offer a lightweight approach to certification by detecting entanglement from one or a few observables, without full state reconstruction [20]–[23]. In distributed settings, such witnesses can often be evaluated using only local measurements and classical communication [33]. Recent work has extended this landscape to detect mixed-state entanglement [34], construct data-driven witnesses [35]–[39] and optimize through parameterized quantum circuits [40]. However, most of them are aimed at detecting whether entanglement is *present* rather than providing quantitative bounds on fidelity with respect to a specific target state. For network verification, what is needed is not a yes/no answer but a certified bound on fidelity that can be compared against the threshold required for teleportation to succeed.

Existing GHZ-state verification protocols. Well-developed verification protocols exist for multipartite entangled states such as GHZ states [41]. When restricted to the bipartite case, the EPR pair corresponds to the $n = 2$ instance of a GHZ state, and the resulting protocol yields a lower bound on fidelity from equatorial-basis correlations. This structure is representative of a broader pattern: methods designed for multipartite certification often provide only a lower bound on bipartite fidelity and do not simultaneously yield an upper bound. Having an upper bound is important for the *no-go* side of the decision: a lower bound alone can certify that a pair is good enough to use, but only an upper bound can certify that a pair is definitively too noisy to use. We provide more analysis and comparison of the GHZ-state verification protocol and our proposed technique in Section IV.

CHSH and fidelity certification. The connection between CHSH violation and the structure of two-qubit entangled states has been studied extensively since the original Bell

and CHSH works and the characterization of two-qubit CHSH violation by Horodecki *et al.* [42]. In the self-testing literature, stronger analytic robustness guarantees relating Bell violation to closeness with an ideal target state were later developed, for example by Kaniewski [43], together with further analyses of the geometry and detectability of entangled states by Bell inequalities [44]. Our work does not attempt to improve on these device-independent self-testing results. Instead, we consider a more restricted setting with fixed, trusted local measurements and ask what explicit bounds on the target-state fidelity $F(\rho) = \langle \Phi^+ | \rho | \Phi^+ \rangle$ can be obtained directly from the resulting CHSH statistic, and how those bounds can be packaged into a practical network-level verification workflow.

III. OVERVIEW AND CHSH-BASED VERIFICATION

In this section, we present the overview of our approach and discuss how CHSH measurements can be used for verification. In the next sections, we describe how this verification algorithm can be integrated for verified quantum communication.

A. Problem Definition and Approach Overview

We consider one quantum-network link between Alice and Bob. Before quantum teleportation can begin, Alice and Bob first share a large number of EPRs. These entanglements are distributed as a bipartite state by an entanglement source. Ideally, each pair of entanglement is in the EPR state $|\Phi^+\rangle$, but in practice noise during generation, transmission, and storage degrades the resource.

Our goal is to verify that the entanglement state is high enough for teleportation to be successful (note here that teleportation is possible even if the EPR states are not maximally entangled). However, verifying the entanglement state collapses the state and can no longer be used for teleportation. Instead, our approach is the following: we assume a batch of $\mathcal{C}$ identical and independently distributed pairs, each in the same state ρ as shown in Figure 1. Alice and Bob sacrifice a random subset of $\beta\mathcal{C}$ pairs for verification and reserve the rest of $(1-\beta)\mathcal{C}$ for teleportation.

Formally, the batch of EPR pairs are modeled as i.i.d.: each pair is an independent copy of the same two-qubit state $\rho \in \mathcal{Q}$, where $\mathcal{Q}$ is the set of 4×4 density matrices. We evaluate this assumption in Section VI. The target state is $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, and the metric that quantifies the quality of ρ is the target-state fidelity $F(\rho) = \langle \Phi^+ | \rho | \Phi^+ \rangle$.

Given a fidelity tolerance α , Alice and Bob want to make the decision

$$\text{go} \quad \text{if } F(\rho) \geq 1 - \alpha, \quad \text{no-go} \quad \text{if } F(\rho) < 1 - \alpha.$$

A *go* decision accepts the batch for teleportation; a *no-go* decision rejects it.

B. Leveraging CHSH Violation

Violation of the CHSH inequality certifies entanglement, but only qualitatively: it rules out local hidden-variable mod-

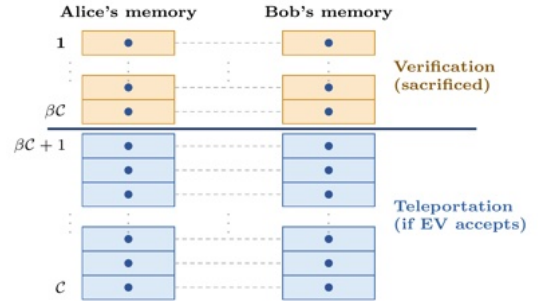


Fig. 1. Batch verification framework. Of $\mathcal{C}$ EPR pairs distributed between Alice and Bob, a fraction $\beta\mathcal{C}$ (orange) are sacrificed for CHSH measurements. The remaining $(1-\beta)\mathcal{C}$ pairs (blue) are reserved for teleportation and released only if the verification test passes; otherwise the entire batch is discarded.

els [45], but does not quantify $F(\rho)$. To make this quantitative, we fix the trusted local observables

$$\begin{aligned} A_0 &= \sigma_x, & A_1 &= \sigma_z, \\ B_0 &= \frac{\sigma_x + \sigma_z}{\sqrt{2}}, & B_1 &= \frac{\sigma_x - \sigma_z}{\sqrt{2}}, \end{aligned} \quad (1)$$

where σ_x and σ_z are the Pauli X and Z matrices, and define the CHSH value

$$S(\rho) = \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle, \quad (2)$$

with $\langle A_i B_j \rangle = \text{Tr}[(A_i \otimes B_j)\rho]$ denoting the expectation of the product of outcomes under settings A_i and B_j . In any local hidden-variable model, $|S| \leq 2$ [1] while Tsirelson's theorem gives $|S| \leq 2\sqrt{2}$ for any quantum state [46]. For $\rho = |\Phi^+\rangle\langle\Phi^+|$, the observables in (1) achieve the maximum $S = 2\sqrt{2}$.

The fact that $|\Phi^+\rangle$ is the unique maximizer of $S(\rho)$ suggests that $S(\rho)$ carries quantitative information about $F(\rho)$: the closer $S(\rho)$ is to $2\sqrt{2}$, the closer ρ should be to the target state. We make this intuition precise in Section IV, where we derive explicit two-sided bounds on $F(\rho)$ in terms of $S(\rho)$.

IV. DERIVING ANALYTIC BOUNDS

Certifying a distributed EPR pair for teleportation requires bounding the target-state fidelity $F(\rho) = \langle \Phi^+ | \rho | \Phi^+ \rangle$ without full state tomography. We use the CHSH value $S(\rho)$ as a single measurable scalar, and establish that it bounds $F(\rho)$ from both sides: a lower bound supporting a *go* decision and an upper bound supporting a *no-go* decision.

Theorem 1. *For any quantum state $\rho \in \mathcal{Q}$, the target-state fidelity satisfies*

$$\frac{S(\rho)}{2\sqrt{2}} \leq F(\rho) \leq \frac{S(\rho)}{4\sqrt{2}} + \frac{1}{2}. \quad (3)$$

Interpretation. We recall the Bell basis $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$, $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. The bounds follow from the representation of the CHSH value

$$S(\rho) = 2\sqrt{2}[p_{\Phi^+} - p_{\Psi^-}], \quad (4)$$

where $p_\phi := \langle \phi | \rho | \phi \rangle$ denotes the Bell-state population. Since $F(\rho) = p_{\Phi^+}$, two elementary constraints on any density

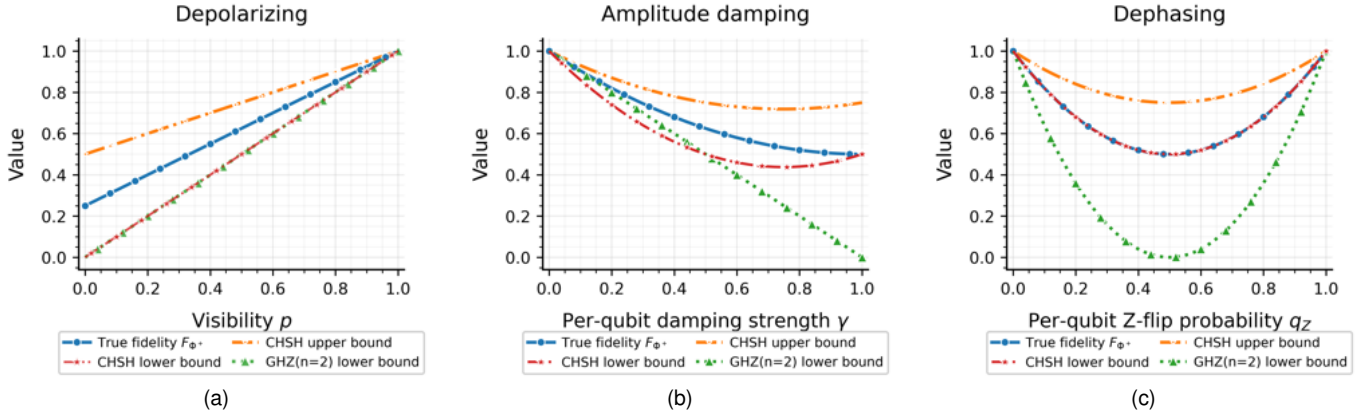


Fig. 2. Comparison of CHSH-based and $n=2$ GHZ-state verification protocols under noise models: (a) depolarizing, (b) amplitude damping, and (c) dephasing.

matrix (positivity and unit trace) are sufficient to bound $F(\rho)$ analytically from $S(\rho)$ alone. A direct consequence is that $S(\rho) = 2\sqrt{2}$ forces $F(\rho) = 1$, i.e. $\rho = |\Phi^+\rangle\langle\Phi^+|$: the CHSH value is uniquely maximized by the target EPR pair.

Tightness of the bounds. The practical value of the bounds depends on how tight they are, which is controlled by different Bell-state populations. The lower bound gap equals exactly $p_{\Psi-}$, while the upper bound gap equals $\frac{1}{2}(p_{\Phi-} + p_{\Psi+})$, corresponding to Y -type and $\{Z, X\}$ -type Pauli errors respectively. In many photonic entanglement sources, $p_{\Psi-}$ can be engineered to be small (for example, in the cluster-state proposal of Lindner and Rudolph [47], Y -type errors from spin precession can be suppressed below 0.2% per photon through low magnetic fields and spectral filtering), making the lower bound physically tight in such settings. The upper bound, by contrast, depends on Z -type and X -type errors, which are not suppressed by the same mechanism. However, it retains its operational value: in the high-fidelity regime the gap $\frac{1}{2}(p_{\Phi-} + p_{\Psi+})$ is automatically small by the unit-trace constraint, so the upper bound gives a tight fidelity estimate; in the low-fidelity regime it still enables a decisive no-go decision regardless of the noise structure. The bound is loosest in the intermediate regime when Z -type and X -type errors are the dominant noise source.

Comparison to Bounds in Existing GHZ Verification. For $n = 2$, the lower bound in the GHZ-state verification protocol [41] reduces to $F(\rho) \geq L_{\text{GHZ2}} := \frac{1}{2}(\langle XX \rangle_\rho - \langle YY \rangle_\rho)$. Two differences arise. First, our protocol also provides an *upper bound*, which the GHZ method lacks. This constrains fidelity overestimation and enables a decisive no-go decision. Second, the two lower bounds are not equally tight under non-isotropic noise. Under depolarizing noise both reduce to $(4F(\rho) - 1)/3$ and coincide (Fig. 2a). Under amplitude damping of per-qubit strength γ , the CHSH bound is $1 - \frac{3}{2}\gamma + \gamma^2$ against $1 - \gamma$ for GHZ, so GHZ is tighter for $\gamma < 1/2$ and CHSH shows an advantage for $\gamma > 1/2$, becoming closer to the fidelity as $\gamma \rightarrow 1$ (Fig. 2b). Under dephasing, the CHSH bound equals $F(\rho)$ exactly while the GHZ bound falls well below it (Fig. 2c).

V. STATISTICAL VERIFICATION FRAMEWORK

We have established the bounds on $F(\rho)$ from $S(\rho)$ that enable certification decisions, but $S(\rho)$ is a quantity that must be estimated from finite measurement outcomes. This section develops the statistical framework to address this problem: we present an estimation algorithm for $S(\rho)$, derive bounds on the fluctuations of the estimator, and use these to construct a hypothesis test that certifies whether $F(\rho)$ meets a prescribed quality threshold.

A. Statistical Estimation for CHSH

Theorem 1 provides a certificate: the closer $S(\rho)$ is to the maximal quantum violation $2\sqrt{2}$, the closer the underlying state is to the EPR pair $|\Phi^+\rangle$. In practice, however, $S(\rho)$ is never directly observed: Alice and Bob have access only to the classical outcomes $a_k, b_k \in \{\pm 1\}$ of local measurements under setting $A_i \otimes B_j$ on their received pairs, and must infer an estimate $\tilde{S}$ from a finite number of those outcomes. This section develops the statistical framework for doing so.

The estimator $\tilde{S}$ in Algorithm 1 is unbiased and its variance decays as $1/N$, meaning the estimate improves as more copies are measured. A practical question is then: how many copies of ρ are needed to estimate $S(\rho)$ within a prescribed additive error ϵ and with a given confidence level $1 - \delta$? We answer this using concentration inequalities, the probabilistic tools that quantify how quickly an empirical average converges to its true value. We present three such guarantees of increasing tightness: a Chebyshev bound (Remark 2) as a simple baseline, a Hoeffding bound (Lemma 3) as the primary worst-case guarantee, and an empirical Bernstein bound (Remark 5) as a sharper data-driven refinement applied after measurement. Corollary 4 then converts the resulting accuracy guarantee on $\tilde{S}$ into a certified interval on $F(\rho)$.

Each outcome product $a_k b_k \in \{\pm 1\}$ is bounded, so the estimator $\tilde{A}_i \tilde{B}_j$ is unbiased with variance

$$\text{Var}[\tilde{A}_i \tilde{B}_j] = \frac{1 - \langle A_i B_j \rangle^2}{N}, \quad (5)$$

Algorithm 1 Estimation of CHSH Measure

- 1: **Input:** Number of samples per setting N (total measurements $4N$).
- 2: Prepare $4N$ copies of ρ .
- 3: **for** each basis pair (A_i, B_j) , $i, j \in \{0, 1\}$ **do**
- 4: Measure N copies, record outcomes (a_k, b_k) for $k = 1, \dots, N$, where $a_k \in \{+1, -1\}$ is Alice's outcome under setting A_i and $b_k \in \{+1, -1\}$ is Bob's outcome under setting B_j , each obtained by local measurement on the k -th shared pair.
- 5: Estimate correlation:

$$\overline{A_i B_j} = \frac{1}{N} \sum_{k=1}^N a_k b_k.$$

6: **end for**

- 7: **Output:** $\bar{S} = \overline{A_0 B_0} + \overline{A_1 B_0} + \overline{A_0 B_1} - \overline{A_1 B_1}$.
-

which decreases with the number of measured copies. Since the four correlators are estimated from disjoint batches, they are statistically independent and their variances add, giving

$$\text{Var}[\bar{S}] = \frac{4 - \sum_{i,j} \langle A_i B_j \rangle^2}{N} \leq \frac{4 - \frac{1}{4}[S(\rho)]^2}{N}. \quad (6)$$

This $1/N$ decay is the structural property that makes concentration inequalities applicable and allows explicit sample complexity bounds to be derived.

Remark 2 (Chebyshev bound). For context, Chebyshev's inequality gives the elementary bound

$$\Pr(|\bar{S} - S(\rho)| > \epsilon) \leq \frac{\text{Var}[\bar{S}]}{\epsilon^2} \leq \frac{4}{N\epsilon^2}.$$

Hence $N = O(1/(\delta\epsilon^2))$ suffices for an ϵ -accurate estimate with confidence $1 - \delta$, improved to $N = 3/(\delta\epsilon^2)$ in the nonlocal regime $S(\rho) \geq 2$. This bound is simple but generally looser than the exponential bounds below.

Lemma 3 (Hoeffding bound). *For $\epsilon \in (0, 2\sqrt{2})$ and $\delta \in (0, 1)$, if*

$$N \geq \frac{32}{\epsilon^2} \ln \left(\frac{2}{1 - (1 - \delta)^{1/4}} \right),$$

then Algorithm 1 produces an estimate $\bar{S}$ satisfying $\Pr(|\bar{S} - S(\rho)| \leq \epsilon) \geq 1 - \delta$.

This follows from a direct application of Hoeffding's inequality to each of the four bounded correlators, combined into a joint confidence statement.

Corollary 4. *Under the sample-size condition in Lemma 3, running Algorithm 1 produces an estimate $\bar{S}$ such that, with probability at least $1 - \delta$,*

$$F(\rho) \in \left[\frac{\bar{S}}{2\sqrt{2}} - \frac{\epsilon}{2\sqrt{2}}, \min \left\{ \frac{\bar{S}}{4\sqrt{2}} + \frac{1}{2} + \frac{\epsilon}{4\sqrt{2}}, 1 \right\} \right].$$

This follows by propagating the CHSH confidence interval from Lemma 3 through the analytical bounds of Theorem 1: a

certified range on $S(\rho)$ translates directly into a certified range on $F(\rho)$.

Remark 5 (Empirical Bernstein refinement). After the N samples are collected, one may tighten the confidence interval post-hoc using an empirical Bernstein bound for each correlator [48]. Let $\hat{V}_{ij} = \frac{1}{N-1} \sum_{k=1}^N (a_k b_k - \overline{A_i B_j})^2$ be the sample variance for setting (i, j) . Then, with probability at least $1 - \delta$,

$$|\bar{S} - S(\rho)| \leq \sum_{i,j \in \{0,1\}} \left[\sqrt{\frac{2\hat{V}_{ij} \ln(8/\delta)}{N}} + \frac{14 \ln(8/\delta)}{3(N-1)} \right].$$

This refinement is especially effective in the high-fidelity regime, where correlator outcomes concentrate near ± 1 and sample variances are small.

Example. For $\delta = 0.1$, $\epsilon = 0.1$, Lemma 3 requires $N \approx 14,000$ per setting ($4N \approx 56,000$ total). Suppose the observed value is $\bar{S} \approx 2.778$; Corollary 4 then yields $F(\rho) \in [0.947, 1]$ with probability 0.90. If sample variances satisfy $\hat{V}_{ij} \approx 0.02$ for all settings, the empirical Bernstein refinement post-hoc shrinks the interval to $F(\rho) \in [0.975, 0.995]$ without any additional measurements.

B. Statistical Verification

The estimation framework above produces a confidence interval on $F(\rho)$, but a network node ultimately needs a binary decision: is this batch of pairs good enough to use, or not? We cast this as a hypothesis test. For a quality tolerance $\alpha \in (0, 1/3)$, the two hypotheses are

$$H_0 : F(\rho) \geq 1 - \alpha \quad \text{vs.} \quad H_1 : F(\rho) \leq 1 - 3\alpha, \quad (7)$$

where accepting H_0 supports the *go* decision (the fidelity is high enough to use the pairs) and rejecting H_0 triggers the *no-go* decision (the fidelity is too low). The gap of 2α between the two thresholds is a deliberate design choice that allows error probabilities to be controlled simultaneously with a finite number of measurements. To implement this test, we propose Algorithm 2, the *CHSH-based Entanglement Verification* (EV) protocol.

Algorithm 2 CHSH-based Entanglement Verification (EV)

- 1: **Input:** $\alpha \in (0, 1/3)$, N
 - 2: Set $\theta_{\text{th}} = 2\sqrt{2} - 5\sqrt{2}\alpha$
 - 3: Run Algorithm 1 with input N and obtain $\bar{S}$
 - 4: **if** $\bar{S} \geq \theta_{\text{th}}$ **then**
 - 5: **Return:** accept H_0
 - 6: **else**
 - 7: **Return:** reject H_0 (equivalently, accept H_1)
 - 8: **end if**
-

Algorithm 2 takes two inputs: the fidelity tolerance α and a sample size N . From α , it computes the decision threshold θ_{th} , calibrated via the analytical bounds of Theorem 1 so that the threshold correctly separates the two fidelity regimes. It then runs Algorithm 1 to obtain $\bar{S}$, and issues a *go* decision if $\bar{S} \geq \theta_{\text{th}}$ and a *no-go* decision otherwise. The following theorem

characterizes its performance guarantees when the number of samples used to estimate $\bar{S}$ is sufficiently large.

Theorem 6 (Entanglement Verification Protocol Performance). *Let*

$$\beta := \sup_{F(\rho) \geq 1-\alpha} \Pr_{\rho}[\bar{S} < \theta_{\text{th}}], \quad \gamma := \sup_{F(\rho) \leq 1-3\alpha} \Pr_{\rho}[\bar{S} \geq \theta_{\text{th}}]$$

denote the type-I and type-II error probabilities of Algorithm 2, respectively. For any $\delta \in (0, 1)$, if the sample size satisfies

$$N \geq \frac{16}{\alpha^2} \ln \left(\frac{2}{1 - (1 - \delta)^{1/4}} \right),$$

then $\beta \leq \delta$ and $\gamma \leq \delta$.

Proof. Write $S = S(\rho)$ and $F = F(\rho)$.

Type-I error. If $F \geq 1 - \alpha$, the upper bound in Theorem 1 gives $S \geq 2\sqrt{2} - 4\sqrt{2}\alpha$, so $\theta_{\text{th}} = 2\sqrt{2} - 5\sqrt{2}\alpha \leq S - \sqrt{2}\alpha$, and $\{\bar{S} < \theta_{\text{th}}\} \subseteq \{|\bar{S} - S| \geq \sqrt{2}\alpha\}$.

Type-II error. If $F \leq 1 - 3\alpha$, the lower bound gives $S \leq 2\sqrt{2} - 6\sqrt{2}\alpha$, so $\theta_{\text{th}} \geq S + \sqrt{2}\alpha$, and $\{\bar{S} \geq \theta_{\text{th}}\} \subseteq \{|\bar{S} - S| \geq \sqrt{2}\alpha\}$.

In both cases, a misclassification requires $|\bar{S} - S| \geq \sqrt{2}\alpha$. Lemma 3 with $\epsilon = \sqrt{2}\alpha$ then gives $\beta, \gamma \leq \delta$ for the prescribed N . $\square$

Comparison to Direct Fidelity Estimation. For the EPR target, DFE reduces to $\hat{F}_{\text{DFE}} = \frac{1}{4}(1 + \langle XX \rangle - \langle YY \rangle + \langle ZZ \rangle)$. Applying DFE to test the same hypothesis in Eq. (7) with symmetric risk δ achieves copy complexity $O(\alpha^{-2} \log(1/\delta))$ [31], matching the asymptotic scaling of the entanglement verification protocol.

Remark 7 (Post-hoc precision via empirical Bernstein). The error bounds hold unconditionally, but their slack depends on the Bell-diagonal structure of ρ : the type-I guarantee is tighter when $\frac{1}{2}(p_{\Phi-} + p_{\Psi+})$ is small, and the type-II guarantee when $p_{\Psi-}$ is small. Both conditions hold in the high-fidelity regime. In this regime the empirical variance $\hat{V}_{ij}$ is also small, and the empirical Bernstein bound (Remark 5) certifies far tighter error probabilities from the same samples. For $\delta = 0.1$ and $\alpha = 0.05$, Hoeffding prescribes $N \approx 28,000$ samples per setting. If $\hat{V}_{ij} \approx 0.02$, the Bernstein bound retroactively certifies $\delta' \approx 6 \times 10^{-23}$, more than twenty orders of magnitude tighter, with no additional measurements.

Limitations. States with fidelity in the indeterminate region $[1 - 3\alpha, 1 - \alpha]$ cannot be classified with the guaranteed error bounds. This gap is a deliberate design choice that trades decisiveness for robustness. Narrowing it requires $N \propto \alpha^{-2}$ samples per setting, so halving the gap quadruples the measurement cost.

This verification procedure is particularly suitable for link-level decisions in quantum networks: it provides an explicit sample complexity of order $O(\alpha^{-2} \log(1/\delta))$ together with worst-case control of both error types. The gap between the worst-case guarantee and typical performance is examined in Section VI-A.

VI. EVALUATION

In this section, we provide an extensive evaluation based on simulating a two-node network using the SeQUeNCe [2] simulator. We integrate our entanglement verification protocol (henceforth called EV). This simulation is able to characterize the performance of EV under realistic network conditions including channel loss, memory decoherence, and storage delay.

The network simulation is complemented by a Monte Carlo study on Werner states. The analytical framework in Section V provides worst-case guarantees, Theorem 6 prescribes a conservative number of pairs per measurement setting N to achieve a given error rate. Our Monte Carlo analysis on Werner states shows that, under depolarizing noise, reliable type-I and type-II error rates are achieved at values of N far below this worst-case prescription. We therefore use values of N suggested by Monte Carlo in the network simulation, where the dominant noise is also depolarizing.

A. Monte Carlo Evaluation on Depolarized States

The sample size per setting N prescribed by Theorem 6 is a worst-case guarantee: it holds for any state $\rho \in \mathcal{Q}$ and any noise realization, without exploiting structure in the noise model. With $\delta = 0.1$, the theorem prescribes $N \approx 2.78 \times 10^4$, 7.72×10^4 , and 6.95×10^5 for $\alpha = 0.05, 0.03$, and 0.01 , respectively, which are substantial costs given limited entanglement generation rates. This raises a natural practical question: *how small can N actually be while still achieving reliable type-I and type-II error rates*, when the distributed state follows a structured noise model rather than the adversarial worst case?

Experiment setting. We performed a Monte Carlo study on Werner states,

$$\rho_W(F) = \frac{4F - 1}{3} |\Phi^+\rangle\langle\Phi^+| + \frac{1 - F}{3} I_4, \quad (8)$$

with fidelity values chosen relative to the entanglement verification decision boundaries $F = 1 - 3\alpha$ (low-fidelity boundary) and $F = 1 - \alpha$ (high-fidelity boundary) for each tested α , as summarized in Table I. For each α , the grid contains three low-fidelity states ($F \leq 1 - 3\alpha$) and three high-fidelity states ($F \geq 1 - \alpha$); states in the ambiguity region $1 - 3\alpha < F < 1 - \alpha$, where our protocol cannot issue a guaranteed decision, are excluded. For each state ρ and sample size N , we generated 10^4 independent trials. In each trial, Algorithm 2 is run with inputs (α, N) to produce the finite-shot estimator $\bar{S}(\rho)$ and apply the decision rule to issue a go/no-go decision.

Performance metrics. For a high-fidelity state ($F \geq 1 - \alpha$), a correct decision is to accept; we report the type-I error rate $\#\{\text{EV rejects } H_0\}/10^4$. For a low-fidelity state ($F \leq 1 - 3\alpha$), a correct decision is to reject; we report the type-II error rate $\#\{\text{EV accepts } H_0\}/10^4$.

Results. Figure 3 shows the empirical type-I and type-II error rates versus N for all tested fidelity levels and values of α , and Table II reports the exact rates at the EV boundary states $F = 1 - 3\alpha$ and $F = 1 - \alpha$. These boundary states are

TABLE I

TESTED FIDELITY VALUES FOR EACH α IN MONTE CARLO SIMULATION, CHOSEN RELATIVE TO THE EV DECISION BOUNDARIES. BOLD VALUES ARE THE EXACT BOUNDARIES $F = 1 - 3\alpha$ (LOW) AND $F = 1 - \alpha$ (HIGH); THE REMAINING VALUES ARE POINTS STRICTLY INSIDE EACH REGION.

α	Low-fidelity ($F \leq 1 - 3\alpha$)	High-fidelity ($F \geq 1 - \alpha$)
0.05	0.75, 0.80, 0.85	0.95 , 0.97, 0.99
0.03	0.81, 0.86, 0.91	0.97 , 0.98, 0.99
0.01	0.91, 0.94, 0.97	0.99 , 0.992, 0.995

the hardest to certify: they lie closest to the decision threshold and require the largest N to achieve small error rates, so we focus on them. The results confirm that the sample requirement of Theorem 6 is substantially conservative, while preserving the theoretical scaling that smaller α requires larger N .

For $\alpha = 0.05$, the type-II error at the low boundary $F = 0.85$ reaches 2×10^{-4} at $N = 750$, and the type-I error at the high boundary $F = 0.95$ reaches 2×10^{-4} at $N = 1000$. Both reach zero by $N = 2000$, roughly $14\times$ fewer samples than the theorem prescribes. For $\alpha = 0.03$, the type-II error at the low boundary $F = 0.91$ reaches zero and the type-I error at the high boundary $F = 0.97$ falls to 7×10^{-4} by $N = 2000$, far below the prescribed $N = 7.72 \times 10^4$. For $\alpha = 0.01$, convergence is slower: at $N = 8000$, the type-II error at the low boundary $F = 0.97$ is 4.8×10^{-3} and the type-I error at the high boundary $F = 0.99$ is 1.73×10^{-2} , both decreasing steadily from 0.2598 and 0.3041 at $N = 500$.

B. Evaluations using a Realistic Network Simulation

We simulate EV as a verification step before quantum teleportation using SeQUeNCe [2], a discrete-event quantum network simulator widely used for realistic modeling of quantum network protocols. The goal is to evaluate how well EV controls type-I and type-II errors under realistic network conditions, and whether pairs that pass verification remain of sufficient quality for teleportation. Whereas the Monte Carlo study evaluates EV on static Werner states with controlled fidelity, the network simulation evaluates the same verification protocol under realistic quantum network conditions including channel loss, storage delay, and memory decoherence. We conduct four experiments (Sections VI-B1–VI-B4), each varying one parameter while holding the others at the baseline values in Table III: the source fidelity F_{src} , the link distance L , the memory coherence time T_m , the sample size per setting N .

Metrics and experimental setup. In our experiments, Alice and Bob share a batch of C EPRs. We implement Algorithm 1 where $4N$ pairs are sacrificed to verify the remaining $C - 4N$ pairs. If EV accepts, those pairs are used for teleportation; otherwise the entire batch is discarded (Fig. 4).

When the EV algorithm is run, we record its accept/reject decision and the mean fidelity of the batch at that moment, denoted F_{ver} . Comparing the decision against the thresholds $1 - \alpha$ and $1 - 3\alpha$ gives the type-I and type-II error rates. We also track the fidelity of the $C - 4N$ pairs reserved for teleportation, denoted F_{rem} , since these pairs continue to de-

cohere while Alice and Bob wait for the EV decision. Finally, under depolarizing noise, we convert F_{rem} to the teleportation fidelity F_{tel} via the relation $F_{\text{tel}} = (1 + 2F_{\text{rem}})/3$ [49].

Each experiment (with different parameter setting) is run 100 times. For batches in the high-quality region ($F_{\text{ver}} \geq 0.97$), a correct decision is to accept; we report the type-I error rate $\#\{\text{EV rejects}\}/100$. For batches in the low-quality region ($F_{\text{ver}} \leq 0.91$), a correct decision is to reject; we report the type-II error rate $\#\{\text{EV accepts}\}/100$. Ambiguous batches ($0.91 < F_{\text{ver}} < 0.97$) are excluded from both calculations and there is no error report for them. The type-II error is the primary safety concern: a falsely accepted low-quality batch caps teleportation fidelity at $F_{\text{tel}} \leq 0.94$.

Baseline parameters. Table III lists the full baseline configuration. We set $\alpha = 0.03$ and $N = 1000$; the Monte Carlo study shows that $N = 1000$ achieves near-zero error rates at tolerance $\alpha = 0.03$, making it a reasonable operating point. With batch size $C = 10000$, a fraction $\beta = 4N/C = 0.4$ of pairs ($4N = 4000$) are sacrificed for verification, leaving 6000 pairs for teleportation. The memory coherence time is $T_m = 100$ ms, representative of coherence times demonstrated in quantum memory experiments and anticipated for future quantum network [50], [51]. Fiber loss is set to $\mu = 0.2$ dB/km, the standard attenuation of single-mode telecom fiber. Channel depolarization is modeled by an exponential decay in fidelity with distance,

$$F_{\text{link}}(L) = F_{\text{min}} + (F_{\text{src}} - F_{\text{min}})e^{-R_c L},$$

where $F_{\text{min}} = 0.25$ is the fidelity of the fully mixed state. These parameters are not tied to a specific hardware platform; they are chosen to cover a realistic operating range.

1) Effect of Source Noise: We vary the source fidelity F to study the performance of EV under different levels of source noise. We test $F \in \{0.90, 0.91, 0.92, 0.93, 0.95, 0.97, 0.98, 0.99\}$.

Fig. 5(a) and Table IV show that type-II error is zero for all low-quality cases: for $F \in \{0.90, 0.91\}$, EV rejects all batches. For $F \in \{0.92, 0.93, 0.95, 0.97\}$, network degradation moves F_{ver} into the ambiguous region, so these points are excluded from the type-I and type-II error calculations. The acceptance rate increases across this transition region, from 0 at $F = 0.92$ to 0.76 at $F = 0.97$. This behavior should not be interpreted as a certified success or failure, because the EV guarantee applies only when $F_{\text{ver}} \leq 1 - 3\alpha$ or $F_{\text{ver}} \geq 1 - \alpha$; it does not apply in the ambiguous region. At $F = 0.98$, the type-I error rate is 7%, and falls further to 2% at $F = 0.99$. At $F = 0.98$, $F_{\text{ver}} = 0.971$ certifies the batches with high quality, and although the remaining pairs decohere slightly to $F_{\text{rem}} = 0.966$ before teleportation, this still yields $F_{\text{tel}} = 0.977$, above $1 - \alpha = 0.97$, confirming that EV-certified batches deliver high-quality teleportation even after accounting for post-verification decoherence.

Overall, EV behaves as a quality-sensitive filter: it rejects every low-quality batch with zero type-II error, accepts high-quality batches reliably, and in the ambiguous region the acceptance rate increases with batch quality. At $F = 0.97$,

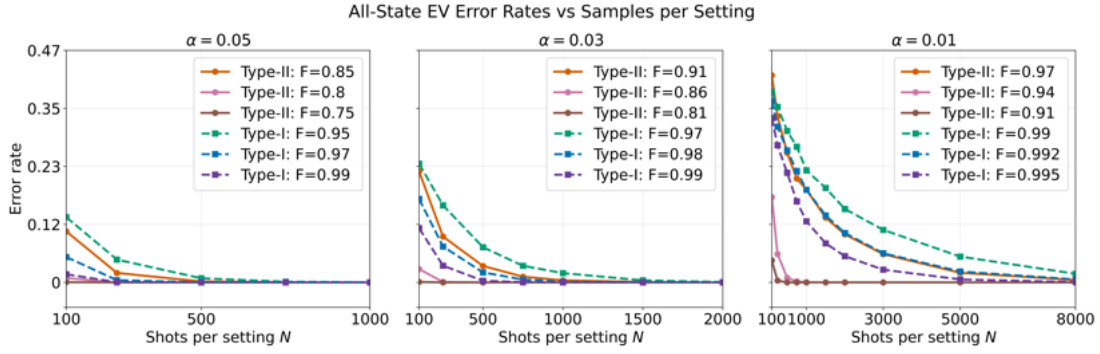


Fig. 3. Empirical EV type-I and type-II error rates versus N for depolarized Werner states, for all tested fidelity levels and $\alpha \in \{0.05, 0.03, 0.01\}$. Solid curves (circular markers): type-II error at low-fidelity states; dashed curves (square markers): type-I error at high-fidelity states. Boundary states at $F = 1 - \alpha$ or $F = 1 - 3\alpha$ exhibit the slowest decay.

TABLE II

EMPIRICAL TYPE-I AND TYPE-II ERROR RATES AT THE EV BOUNDARY STATES, ESTIMATED FROM 10^4 MONTE CARLO TRIALS ON DEPOLARIZED WERNER STATES. FOR EACH α , THE LOW-FIDELITY BOUNDARY $F = 1 - 3\alpha$ REPORTS THE TYPE-II ERROR (FALSE ACCEPTANCE RATE) AND THE HIGH-FIDELITY BOUNDARY $F = 1 - \alpha$ REPORTS THE TYPE-I ERROR (FALSE REJECTION RATE). THESE BOUNDARY STATES ARE THE HARDEST TO CERTIFY AND SET THE EFFECTIVE SAMPLE-SIZE REQUIREMENT. DASHES INDICATE SAMPLE SIZES NOT EVALUATED FOR THAT α .

α	Boundary	Error type	$N = 500$	$N = 750$	$N = 1000$	$N = 1500$	$N = 2000$	$N = 5000$	$N = 8000$
0.05	$F = 1 - 3\alpha = 0.85$	Type-II	1.9×10^{-3}	2×10^{-4}	0	0	0	—	—
	$F = 1 - \alpha = 0.95$	Type-I	8.3×10^{-3}	1.4×10^{-3}	2×10^{-4}	1×10^{-4}	0	—	—
0.03	$F = 1 - 3\alpha = 0.91$	Type-II	3.3×10^{-2}	1.15×10^{-2}	4.2×10^{-3}	8×10^{-4}	0	—	—
	$F = 1 - \alpha = 0.97$	Type-I	7.06×10^{-2}	3.33×10^{-2}	1.86×10^{-2}	4.2×10^{-3}	7×10^{-4}	—	—
0.01	$F = 1 - 3\alpha = 0.97$	Type-II	0.2598	0.2084	0.1860	0.1300	0.0964	0.0193	0.0048
	$F = 1 - \alpha = 0.99$	Type-I	0.3041	0.2722	0.2247	0.1894	0.1473	0.0518	0.0173

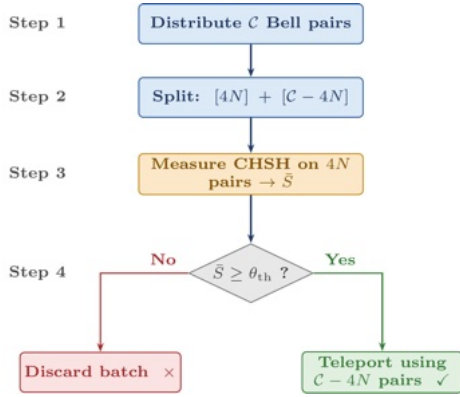


Fig. 4. EV-gated teleportation protocol. Of C Bell pairs distributed between Alice and Bob, $4N$ are measured to estimate the CHSH statistic $\tilde{S}$. If $\tilde{S} \geq \theta_{th}$, EV issues a go decision and the remaining $C - 4N$ pairs are used for teleportation; otherwise the batch is discarded.

76% of batches are accepted with $F_{tel} = 0.971$, exceeding $1 - \alpha = 0.97$ despite falling outside the theoretically guaranteed region, demonstrating empirically reliable behavior. The protocol thus remains practically useful across the full range of source noise.

2) *Effect of Link Distance*: In this experiment, we vary the distance between Alice and Bob, $L \in \{1, 4, 8, 12, 20, 30, 40\}$ km. As link distance increases, higher transmission loss means more attempts are needed

TABLE III

BASELINE PARAMETERS FOR THE EV PROTOCOL EVALUATION IN THE NETWORK SIMULATION. THE FIRST BLOCK IS FIXED ACROSS ALL EXPERIMENTS; THE SECOND BLOCK LISTS PARAMETERS THAT ARE VARIED INDIVIDUALLY, WITH THEIR DEFAULT VALUES.

Parameter	Symbol	Value
<i>Fixed across all experiments</i>		
EV precision	α	0.03
Batch size	C	10000
Fiber attenuation	μ	0.2 dB/km
Channel depolarizing rate	R_c	0.002 km^{-1}
Fidelity floor	F_{min}	0.25
Source emission period	τ_{src}	$0.1 \mu\text{s}$
CHSH measurement time per pair	τ_{meas}	$0.1 \mu\text{s}$
<i>Changed in specific experiments</i>		
# of measurements per setting	N	1000
Source Bell-pair fidelity	F	1.0
Node separation	L	1 km
Memory coherence time	T_m	100 ms

to accumulate a full batch of C pairs, increasing the time required to complete each verification round, subject to memory decoherence. At the same time, longer channels introduce more depolarization noise, also degrading the quality of the accumulated pairs. This experiment therefore studies how these effects of distance jointly impact verification performance.

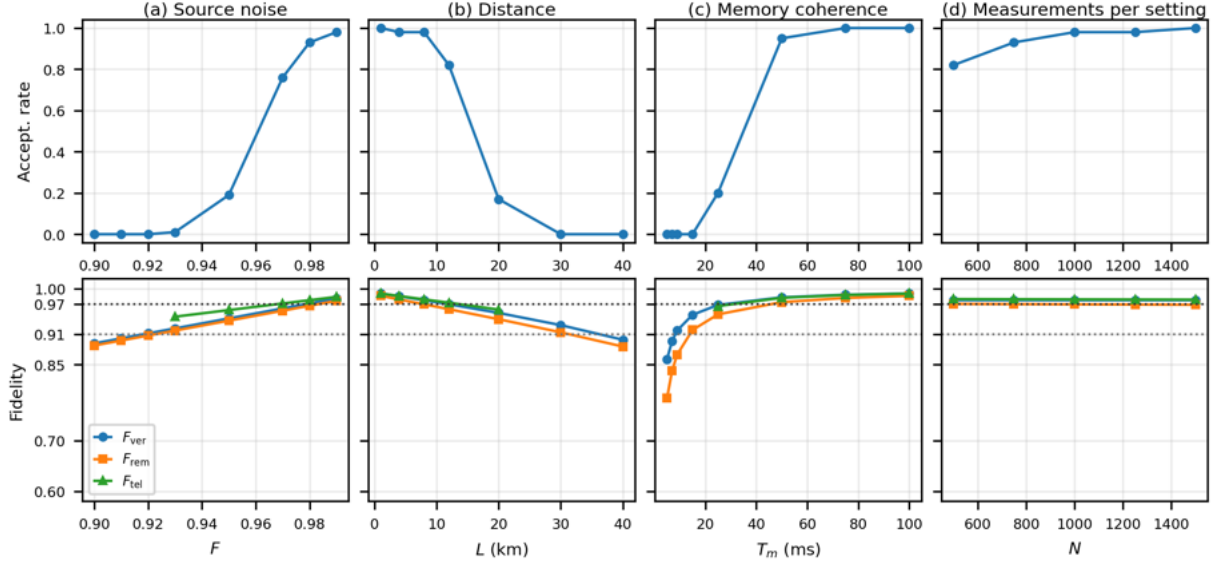


Fig. 5. Network simulation results for EV before teleportation with $\alpha = 0.03$ and $N = 1000$, except in panel (d), where N is varied. Each column varies one parameter while keeping the others fixed: (a) source fidelity F , (b) link distance L , (c) memory coherence time T_m , and (d) number of CHSH measurements per setting N , with $L = 8$ km. The top row shows the EV acceptance rate. The bottom row shows the verification-time fidelity F_{ver} , the remaining-pair fidelity F_{rem} , and the teleportation fidelity F_{tel} . Dotted horizontal lines mark the EV decision boundaries $1 - 3\alpha = 0.91$ and $1 - \alpha = 0.97$. Points between these two boundaries lie in the ambiguous region, where type-I and type-II error guarantees are not reported.

TABLE IV

EFFECT OF SOURCE NOISE ON EV PERFORMANCE ($\alpha = 0.03$, $N = 1000$, AVERAGED OVER 100 BATCHES). THE EV ERROR COLUMN REPORTS TYPE-II ERROR IN THE LOW REGION AND TYPE-I ERROR IN THE HIGH REGION; IT IS NOT REPORTED FOR AMBIG. POINTS. “—” INDICATES NOT APPLICABLE.

F	Region	Accept	EV error	F_{ver}	F_{rem}	F_{tel}
0.90	Low	0.00	0.00	0.892	0.888	—
0.91	Low	0.00	0.00	0.902	0.898	—
0.92	Ambig.	0.00	—	0.912	0.907	—
0.93	Ambig.	0.01	—	0.922	0.917	0.945
0.95	Ambig.	0.19	—	0.941	0.937	0.958
0.97	Ambig.	0.76	—	0.961	0.956	0.971
0.98	High	0.93	0.07	0.971	0.966	0.977
0.99	High	0.98	0.02	0.981	0.976	0.984

Fig. 5(b) and Table V show that EV maintains correct behavior across all tested distances, with performance characteristics that depend on which fidelity region the link operates in. At short distances ($L \leq 8$ km), batches fall well within the High region and EV performs reliably, with type-I error at most 0.02 and high acceptance rate. At intermediate distances ($L = 12$ – 30 km), batches enter the ambiguous region and EV withholds certification. The acceptance rate falls from 0.82 at $L = 12$ km to zero by $L = 30$ km, consistent with increased channel depolarization and memory decoherence over longer accumulation times pushing CHSH statistics below the acceptance threshold. At $L = 40$ km, batches fall into the Low region and EV rejects all of them with zero type-II error.

In the intermediate region, the acceptance rate increases with batch quality, so accepted batches are more likely to have

TABLE V

EFFECT OF LINK DISTANCE ON EV PERFORMANCE ($\alpha = 0.03$, $N = 1000$, AVERAGED OVER 100 BATCHES). THE EV ERROR COLUMN REPORTS TYPE-II ERROR IN THE LOW REGION AND TYPE-I ERROR IN THE HIGH REGION; IT IS NOT REPORTED FOR AMBIG. POINTS. “—” INDICATES NOT APPLICABLE.

L (km)	Region	Accept	EV error	F_{ver}	F_{rem}	F_{tel}
1	High	1.00	0.00	0.991	0.986	0.990
4	High	0.98	0.02	0.985	0.979	0.986
8	High	0.98	0.02	0.977	0.969	0.979
12	Ambig.	0.82	—	0.969	0.959	0.972
20	Ambig.	0.17	—	0.952	0.940	0.959
30	Ambig.	0.00	—	0.928	0.914	—
40	Low	0.00	0.00	0.899	0.886	—

good quality. At $L = 12$ km, for instance, 82% of batches are accepted with $F_{\text{tel}} = 0.972$, exceeding $1 - \alpha = 0.97$ despite falling outside the formally guaranteed region.

3) *Effect of Memory Coherence*: We vary the memory coherence time $T_m \in \{5, 7, 9, 15, 25, 50, 75, 100\}$ ms.

Shorter coherence times cause pairs to decohere faster while waiting in memory during batch accumulation, directly reducing their fidelity before verification even begins. This experiment therefore isolates the effect of memory quality on EV performance, independent of channel loss or source imperfections.

Fig. 5(c) and Table VI show a clear threshold behavior as T_m increases. At $T_m \leq 7$ ms, rapid decoherence pushes F_{ver} into the Low region. EV achieves zero type-II error, correctly rejecting all low-quality batches without ever falsely certifying them. For $T_m \in \{9, 15, 25\}$ ms, batches fall in the ambiguous

TABLE VI

EFFECT OF MEMORY COHERENCE TIME ON EV PERFORMANCE ($\alpha = 0.03$, $N = 1000$, AVERAGED OVER 100 BATCHES). THE EV ERROR COLUMN REPORTS TYPE-II ERROR IN THE LOW REGION AND TYPE-I ERROR IN THE HIGH REGION; IT IS NOT REPORTED FOR AMBIG. POINTS. “—” INDICATES NOT APPLICABLE.

T_m (ms)	Region	Accept	EV error	F_{ver}	F_{rem}	F_{tel}
5	Low	0.00	0.00	0.861	0.784	—
7	Low	0.00	0.00	0.896	0.838	—
9	Ambig.	0.00	—	0.917	0.870	—
15	Ambig.	0.00	—	0.948	0.918	—
25	Ambig.	0.20	—	0.968	0.949	0.966
50	High	0.95	0.05	0.983	0.973	0.982
75	High	1.00	0.00	0.988	0.982	0.988
100	High	1.00	0.00	0.991	0.986	0.990

region and type-I/II errors are not defined. The acceptance rate rises from 0.00 to 0.20 across this range, but EV provides no certification guarantee here. At $T_m \geq 50$ ms, batches enter the High region. Type-I error is 0.05 at $T_m = 50$ ms, meaning EV occasionally rejects high-quality batches due to statistical fluctuations in the CHSH estimator near the boundary with $F_{\text{rem}} = 0.973$. At $T_m \geq 75$ ms, type-I error reaches 0.00 and EV accepts all batches.

4) *Effect of Number of Measurements*: In the static Monte Carlo setting, increasing the number of CHSH measurements per setting N improves the accuracy of EV. In the network setting, however, larger N also has operational costs: EV consumes $4N$ pairs, leaving fewer pairs for teleportation, and it takes longer to perform the measurements, letting the remaining pairs decohere longer in memory. Thus, varying N exposes a three-way tradeoff between verification accuracy, the number of pairs left for teleportation, and the fidelity of those remaining pairs. We fix $L = 8$ km, where F_{ver} sits close to the High-region boundary, making type-I error most sensitive to N , and vary $N \in \{500, 750, 1000, 1250, 1500\}$, corresponding to verification fractions $\beta = 4N/C$ ranging from 0.20 to 0.60.

Fig. 5(d) and Table VII show that all tested values of N remain in the High region with $F_{\text{ver}} \approx 0.977$. The dominant effect of increasing N is a reduction in type-I error: it falls from 0.18 at $N = 500$ to 0.00 at $N = 1500$, while the acceptance rate rises from 0.82 to 1.00.

The cost of larger N is a marginal decrease in remaining-pair quality. F_{rem} decreases from 0.970 at $N = 500$ to 0.968 at $N = 1500$, and F_{tel} from 0.979 to 0.978, remaining well above the $1 - \alpha = 0.97$ throughout. While small in absolute terms, this illustrates the network-level tradeoff absent in static analysis: more measurements reduce type-I error but extend the decoherence window for the unverified pairs.

Finding. Across the four network experiments, the type-II error rate is zero for every low-quality configuration tested, confirming that EV never commits low-fidelity pairs to teleportation regardless of the source of degradation. In the high-quality regime, EV accepts batches reliably with high acceptance rate. In the ambiguous region, the acceptance rate increases with batch quality, so accepted batches are more

TABLE VII

EFFECT OF THE NUMBER OF CHSH MEASUREMENTS PER SETTING ON EV PERFORMANCE ($\alpha = 0.03$, $L = 8$ km, AVERAGED OVER 100 BATCHES). THE EV ERROR COLUMN REPORTS TYPE-I ERROR BECAUSE ALL TESTED POINTS ARE IN THE HIGH REGION.

N	Region	Accept	EV error	F_{ver}	F_{rem}	F_{tel}
500	High	0.82	0.18	0.977	0.970	0.979
750	High	0.93	0.07	0.977	0.970	0.979
1000	High	0.98	0.02	0.977	0.969	0.979
1250	High	0.98	0.02	0.977	0.969	0.979
1500	High	1.00	0.00	0.977	0.968	0.978

likely to carry high fidelity; the network simulations consistently show F_{tel} well above useful thresholds $1 - \alpha = 0.97$ for accepted batches even under realistic channel loss and memory decoherence. The gap between F_{rem} and F_{ver} due to post-verification decoherence is modest in practice and does not undermine teleportation quality F_{tel} . These results confirm that EV remains a safe and practically useful primitive across the full range of network conditions tested.

On the i.i.d. assumption. Our analysis treats the batch as i.i.d. copies of one state ρ , which keeps the analysis simple. In practice the channel can change while the batch is collected, for example as the memory loses coherence, so pairs delivered at different times can be in slightly different states, and noise can affect several pairs together and make them correlated. Because the tested pairs are chosen at random, they are a fair sample of the whole batch, so the go/no-go decision stays meaningful even when the pairs are not perfectly identical. The results above support this: across the network experiments the protocol rejected every low-fidelity batch, even under realistic channel loss and memory decoherence. Extending the analysis to correlated or non-identical pairs is a natural next step [52].

VII. CONCLUSION

We have developed a CHSH-based entanglement verification framework connecting three layers: analytical bounds relating $S(\rho)$ to $F(\rho)$, enabling both go and no-go decisions from a single observable; a statistical EV protocol with explicit sample complexity $O(\alpha^{-2} \log(1/\delta))$ and worst-case type-I/type-II error guarantees; and Monte Carlo validation showing that the worst-case bound is substantially conservative under depolarized noise, with reliable performance achieved at far smaller number of measurement than the bound prescribes. Network simulations using SeQuENCe show that the type-II error rate is zero across all tested low-quality configurations, consistent with the safety guarantee. Physical-layer degradation manifests primarily as increased type-I error rather than unsafe certification, and the post-verification fidelity drop is absorbed by the teleportation fidelity relation, preserving downstream quality. These results position CHSH-based EV as an analytically grounded primitive for link-level certification. Open directions include extension to multi-hop entanglement distribution with swapping, integration with purification protocols, and tighter complexity bounds under structured noise models beyond depolarizing noise.

REFERENCES

- [1] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, Oct 1969. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.23.880>
- [2] X. Wu, A. Kolar, J. Chung, D. Jin, T. Zhong, M. Suchara, and R. Kettimuthu, "SeQUeNCe: a customizable discrete-event simulator of quantum networks," *Quantum Science and Technology*, vol. 6, no. 4, p. 045027, 2021.
- [3] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, 1984, pp. 175–179. [Online]. Available: <https://doi.org/10.1016/j.tcs.2014.05.025>
- [4] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug 1991. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.67.661>
- [5] J. F. Fitzsimons, "Private quantum computation: an introduction to blind quantum computing and related protocols," *npj Quantum Information*, vol. 3, p. 23, 2017. [Online]. Available: <https://www.nature.com/articles/s41534-017-0025-3>
- [6] S. Barz, E. Kashefi, A. Broadbent, J. F. Fitzsimons, A. Zeilinger, and P. Walther, "Demonstration of blind quantum computing," *Science*, vol. 335, no. 6066, pp. 303–308, 2012. [Online]. Available: <https://www.science.org/doi/10.1126/science.1214707>
- [7] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar 1993. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.70.1895>
- [8] S. L. N. Hermans, M. Pompili, H. K. C. Beukers *et al.*, "Qubit teleportation between non-neighbouring nodes in a quantum network," *Nature*, vol. 605, pp. 663–668, 2022. [Online]. Available: <https://www.nature.com/articles/s41586-022-04697-y>
- [9] M. Caleffi, A. S. Cacciapuoti, and G. Bianchi, "Quantum internet: from communication to distributed computing!" in *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication*, ser. NANOCOM '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: <https://doi.org/10.1145/3233188.3233224>
- [10] S. Muralidharan, L. Li, J. Kim *et al.*, "Optimal architectures for long distance quantum communication," *Scientific Reports*, vol. 6, p. 20463, 2016. [Online]. Available: <https://doi.org/10.1038/srep20463>
- [11] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum repeaters: The role of imperfect local operations in quantum communication," *Phys. Rev. Lett.*, vol. 81, no. 26, pp. 5932–5935, Dec 1998. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.81.5932>
- [12] J. Z. Bernád, "Hybrid quantum repeater based on resonant qubit-field interactions," *Phys. Rev. A*, vol. 96, p. 052329, Nov 2017. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.96.052329>
- [13] W. Munro, A. Stephens, S. Devitt *et al.*, "Quantum communication without the necessity of quantum memories," *Nature Photonics*, vol. 6, pp. 777–781, 2012. [Online]. Available: <https://doi.org/10.1038/nphoton.2012.243>
- [14] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of noisy entanglement and faithful teleportation via noisy channels," *Phys. Rev. Lett.*, vol. 76, pp. 722–725, Jan 1996. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.76.722>
- [15] M. Vitoria, S. Tserkis, S. Krastanov, A. S. de la Cerda, S. Willis, and P. Narang, "Entanglement purification on quantum networks," *Phys. Rev. Res.*, vol. 5, p. 033171, Sep 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevResearch.5.033171>
- [16] S. Pallister, N. Linden, and A. Montanaro, "Optimal verification of entangled states with local measurements," *Phys. Rev. Lett.*, vol. 120, no. 17, p. 170502, 2018.
- [17] R. Blume-Kohout, J. O. Yin, and S. J. van Enk, "Entanglement verification with finite data," *Phys. Rev. Lett.*, vol. 105, p. 170501, 2010.
- [18] D. Lu, T. Xin, N. Yu, Z. Ji, J. Chen, G. Long, J. Baugh, X. Peng, B. Zeng, and R. Laflamme, "Tomography is necessary for universal entanglement detection with single-copy observables," *Phys. Rev. Lett.*, vol. 116, p. 230501, Jun 2016. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.116.230501>
- [19] W.-H. Zhang, C. Zhang *et al.*, "Experimental optimal verification of entangled states using local measurements," *Phys. Rev. Lett.*, vol. 125, p. 030506, 2020.
- [20] O. Gühne and G. Tóth, "Entanglement detection," *Physics Reports*, vol. 474, no. 1–6, pp. 1–75, 2009. [Online]. Available: <https://doi.org/10.1016/j.physrep.2009.02.004>
- [21] A. Bera, J. Bae, B. C. Hiesmayr *et al.*, "On the structure of mirrored operators obtained from optimal entanglement witnesses," *Scientific Reports*, vol. 13, p. 10733, 2023. [Online]. Available: <https://doi.org/10.1038/s41598-023-37771-0>
- [22] K. Siudzińska and D. Chruściński, "Entanglement witnesses from mutually unbiased measurements," *Scientific Reports*, vol. 11, p. 22988, 2021. [Online]. Available: <https://www.nature.com/articles/s41598-021-02356-2>
- [23] D. Chruściński and G. Sarbicki, "Entanglement witnesses: construction, analysis and classification," *J. Phys. A: Math. Theor.*, vol. 47, no. 48, p. 483001, 2014. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1751-8113/47/48/483001>
- [24] D. Aharonov, A. W. Harrow, Z. Landau, D. Nagaj, M. Szegedy, and U. Vazirani, "Local tests of global entanglement and a counterexample to the generalized area law," in *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*. IEEE, 2014, pp. 246–255.
- [25] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, "Sample-optimal tomography of quantum states," in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, ser. STOC '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 913–925. [Online]. Available: <https://doi.org/10.1145/2897518.2897585>
- [26] T. Xin, D. Lu, J. Klassen, N. Yu, Z. Ji, J. Chen, X. Ma, G. Long, B. Zeng, and R. Laflamme, "Quantum state tomography via reduced density matrices," *Phys. Rev. Lett.*, vol. 118, p. 020401, Jan 2017. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.118.020401>
- [27] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne, "Optimizing shadow tomography with generalized measurements," *Phys. Rev. Lett.*, vol. 129, p. 220502, Nov 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.129.220502>
- [28] V. Zapatero, T. van Leent, R. Arnon-Friedman *et al.*, "Advances in device-independent quantum key distribution," *npj Quantum Information*, vol. 9, p. 10, 2023. [Online]. Available: <https://www.nature.com/articles/s41534-023-00684-x>
- [29] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C. C.-W. Lim *et al.*, "A device-independent quantum key distribution system for distant users," *Nature*, vol. 607, no. 7920, pp. 687–691, 2022.
- [30] Y. Zeng, Z. Wu, X. Du Trinh, Y. Yang, N. Yu, and A. Balasubramanian, "Rethinking quantum network design using a verification-based quantum transmission protocol," in *IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2026.
- [31] S. T. Flammia and Y.-K. Liu, "Direct fidelity estimation from few Pauli measurements," *Phys. Rev. Lett.*, vol. 106, no. 23, p. 230501, 2011.
- [32] J. Barberà-Rodríguez, M. Navarro, and L. Zambrano, "Sampling groups of Pauli operators to enhance direct fidelity estimation," *Quantum*, vol. 9, p. 1784, 2025.
- [33] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera, "Detection of entanglement with few local measurements," *Phys. Rev. A*, vol. 66, p. 062305, Dec 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.66.062305>
- [34] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmon, P. Calabrese, B. Kraus, J. Preskill, P. Zoller, and B. Vermersch, "Mixed-state entanglement from local randomized measurements," *Phys. Rev. Lett.*, vol. 125, p. 200501, 2020.
- [35] J. Roik, K. Bartkiewicz, A. Černoč, and K. Lemr, "Accuracy of entanglement detection via artificial neural networks and human-designed entanglement witnesses," *Phys. Rev. Appl.*, vol. 15, p. 054006, 2021.
- [36] D. Koutný, L. Ginés, M. Moczala-Dusanowska, S. Höfling, C. Schneider, A. Predojević, and M. Ježek, "Deep learning of quantum entanglement from incomplete measurements," *Science Advances*, vol. 9, no. 29, p. eadd7131, 2023. [Online]. Available: <https://www.science.org/doi/abs/10.1126/sciadv.add7131>
- [37] N. Asif, U. Khalid, A. Khan *et al.*, "Entanglement detection with artificial neural networks," *Scientific Reports*, vol. 13, p. 1562, 2023. [Online]. Available: <https://doi.org/10.1038/s41598-023-28745-3>
- [38] J. Ureña, A. Sojo, J. Bermejo-Vega *et al.*, "Entanglement detection with classical deep neural networks," *Scientific Reports*, vol. 14, p. 18109, 2024. [Online]. Available: <https://doi.org/10.1038/s41598-024-68213-0>

- [39] A. C. Greenwood, L. T. Wu, E. Y. Zhu, B. T. Kirby, and L. Qian, “Machine-learning-derived entanglement witnesses,” *Phys. Rev. Appl.*, vol. 19, p. 034058, Mar 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevApplied.19.034058>
- [40] H. Matsunaga and L. B. Ho, “Detecting and protecting entanglement through nonlocality, variational entanglement witness, and nonlocal measurements,” *Phys. Rev. Res.*, vol. 7, p. 013239, 2025.
- [41] W. McCutcheon, A. Pappa, B. A. Bell, A. McMillan, A. Chailloux, T. Lawson, M. Mafu, D. Markham, E. Diamanti, I. Kerenidis, J. G. Rarity, and M. S. Tame, “Experimental verification of multipartite entanglement in quantum networks,” *Nature Communications*, vol. 7, p. 13251, 2016. [Online]. Available: <https://www.nature.com/articles/ncomms13251>
- [42] R. Horodecki, P. Horodecki, and M. Horodecki, “Violating Bell inequality by mixed spin- $\frac{1}{2}$ states: necessary and sufficient condition,” *Physics Letters A*, vol. 200, no. 5, pp. 340–344, 1995. [Online]. Available: [https://doi.org/10.1016/0375-9601\(95\)00214-N](https://doi.org/10.1016/0375-9601(95)00214-N)
- [43] J. Kaniewski, “Analytic and nearly optimal self-testing bounds for the Clauser-Horne-Shimony-Holt and Mermin inequalities,” *Phys. Rev. Lett.*, vol. 117, no. 7, p. 070402, 2016. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.117.070402>
- [44] A. Sauer, J. Z. Bernád, H. J. Moreno, and G. Alber, “Entanglement in bipartite quantum systems: Euclidean volume ratios and detectability by Bell inequalities,” *J. Phys. A: Math. Theor.*, vol. 54, no. 49, p. 495302, 2021.
- [45] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*. Cambridge, England: Cambridge University Press, 1987.
- [46] B. S. Cirel’son, “Quantum generalizations of Bell’s inequality,” *Letters in Mathematical Physics*, vol. 4, pp. 93–100, 1980. [Online]. Available: <https://doi.org/10.1007/BF00417500>
- [47] N. H. Lindner and T. Rudolph, “Proposal for pulsed on-demand sources of photonic cluster state strings,” *Phys. Rev. Lett.*, vol. 103, p. 113602, 2009.
- [48] A. Maurer and M. Pontil, “Empirical Bernstein bounds and sample variance penalization,” in *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*, 2009.
- [49] M. Horodecki, P. Horodecki, and R. Horodecki, “General teleportation channel, singlet fraction, and quasidistillation,” *Phys. Rev. A*, vol. 60, pp. 1888–1898, Sep 1999. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.60.1888>
- [50] M. Körber, O. Morin, S. Langenfeld, A. Neuzner, S. Ritter, and G. Rempe, “Decoherence-protected memory for a single-photon qubit,” *Nature Photonics*, vol. 12, pp. 18–21, 2018. [Online]. Available: <https://doi.org/10.1038/s41566-017-0050-y>
- [51] W. Tittel, M. Afzelius, A. Kinos, L. Rippe, and A. Walther, “Quantum networks using rare-earth ions,” *Quantum Science and Technology*, vol. 10, p. 033002, 2025. [Online]. Available: <https://iopscience.iop.org/article/10.1088/2058-9565/add93>
- [52] N. Yu, “What if you have only one copy? low-depth quantum circuits have no advantage in decision problems!” 2024. [Online]. Available: <https://arxiv.org/abs/2404.05714>