Compressed Sensing for Efficient Fidelity Estimation of GHZ States

Farrokh Labib,[1,*] David Nicholaeff,[2] Vincent Russo,[1] and William Zeng[1,3]

[1] *Unitary Foundation*
[2] *New Mexico Consortium*
[3] *Quantonation*

(Dated: April 23, 2026)

Accurately characterizing multipartite entangled states is a critical challenge in quantum information processing. In this work, we focus on applying compressed sensing techniques to efficiently estimate the fidelity of Greenberger-Horne-Zeilinger (GHZ) states. By exploiting the inherent sparsity of these states, our compressed sensing protocol drastically reduces the measurement overhead traditionally required for state verification while maintaining high accuracy. To evaluate the practical performance of this approach, we test the protocol on GHZ states using both quantum simulators and Quantinuum's trapped-ion hardware. Furthermore, we implement error detection techniques during our hardware evaluations, demonstrating the robustness and viability of compressed sensing for fidelity estimation in noisy experimental environments.

I. INTRODUCTION

The creation and verification of large-scale entangled states are fundamental tasks in quantum computing technology. Among these, the Greenberger-Horne-Zeilinger (GHZ) state[11]—a maximally entangled state of N qubits—serves as a critical benchmark for quantum hardware performance, vital for applications in quantum secret sharing [15], quantum metrology [10], error correction [30], and many others.

To utilize GHZ states as benchmarks, we have to reliably and efficiently estimate the fidelity of the prepared quantum state. Standard quantum state tomography, or Direct Fidelity Estimation [5, 8], is prohibitively expensive for large N, and while Parity Oscillation methods [13, 23, 32] are more efficient, they still require a large number of data points to satisfy the Nyquist-Shannon sampling theorem. To overcome these limitations, we observe that parity oscillation signals from the quantum measurements are "sparse signals". This allows us to use tools from the field of compressed sensing (CS) [4] to exponentially reduce the number of data points needed to reconstruct the parity oscillation signal and hence the coherence of the prepared GHZ state. We want to note that compressed sensing as a general technique has already been widely used in quantum information, for example, in state tomography [2, 12], Hamiltonian learning [20, 29], and there are many other examples.

To evaluate the practical performance of this compressed sensing approach, we test the protocol on GHZ states using both quantum simulators and actual quantum hardware. However, preparing large GHZ states remains a significant challenge due to their extreme sensitivity to imperfections in the experiment. There have been many attempts at preparing large-scale GHZ states in many different quantum computing platforms like ion-traps [18, 23, 28], neutral atoms [27, 31], photons [36], and superconducting [1, 19, 24, 32, 37]. The most recent and largest-scale GHZ state is due to the recent work by IBM [16], where they prepare a 120-qubit GHZ state with fidelity > 0.5. The main tools used in that work are based on the low-overhead error detection ideas developed in [22] by the same group. The main idea is to use pairs of qubits in the GHZ states to apply a parity check that maximizes "coverage" (this is defined in Section II A).

Building upon these recent low-overhead error detection strategies, our work adapts this framework to the all-to-all connectivity of trapped-ion systems such as Quantinuum hardware. First, this allows us to prepare GHZ states of size N in $\log(N)$ depth, significantly reducing errors. Second, the choice of parity checks becomes significantly more simplified, and we can reach much higher coverage of qubits with fewer parity checks, thus reducing the amount of extra operations needed and thus reducing the potential of adding more noise.

The aim of this paper is twofold: 1) we want to show that there are compressed sensing techniques that can be used to efficiently estimate the fidelity of GHZ states using the parity oscillation method, and 2) that this protocol can be successfully tested on simulators and actual hardware by applying an error-detection protocol using flag qubits, which works particularly well on all-to-all connected devices.

II. PRELIMINARIES

In this section, we lay the groundwork for the methods we use. We begin by describing the structure of GHZ states and the specific error detection strategy using flag qubits. We then outline the theoretical basis for efficient verification of the fidelity: first, by detailing the parity oscillation method for measuring coherence, and second, by introducing the compressed sensing framework that allows us to estimate fidelity efficiently.

* farrokh@unitary.foundation

A. Error detection on GHZ states

The N-qubit GHZ state is a maximally entangled state defined as

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}) \tag{1}$$

It has its stabilizer group generated by Z_iZ_{i+1} for $i = 1, \ldots, n-1$ and $XX\ldots X$. Therefore, we can coherently check the parity of any two qubits i, j by introducing an ancilla qubit and applying CNOTs with control qubits i and j to the ancilla. Post-selecting on the ancillas being in the zero state gives us an error detection protocol. We call these ancilla qubits flag qubits.

In this work, the N-qubit GHZ state is prepared using a binary tree structure of CNOT gates, starting from a root Bell pair and expanding outwards. This logarithmic-depth circuit is efficient but susceptible to errors near the root, which can propagate to affect a large fraction of the system (high-weight errors). Error detection protocols employ flag qubits to detect dangerous faults that could otherwise propagate and corrupt the logical state. In our approach, we leverage the flexibility of the assumed all-to-all connectivity to place these flag qubits strategically within the GHZ state preparation circuit.

To enhance fidelity using error detection and post-selection, we implement a flag strategy utilizing optimized ZZ checks: we employ flag qubits to perform ZZ measurements on subsets of the data qubits at the end of the preparation circuit such that we maximize the "coverage" of that check [16].

The coverage of a check is defined to be the size of the path from the two qubits on which the check is applied to the least common ancestor following the path of the CNOTs backwards in time. By the size of the path, we mean the number of qubits we visit on this path. The coverage ratio is this size divided by the total number of qubits. Intuitively, for a given check, the larger the coverage, the more errors it can potentially detect. Any bitflip error that might happen on a qubit in the path of a check will propagate to the flag qubit. So we want to pick pairs of qubits in the GHZ state such that we maximize the total coverage of all the flag qubit checks. Figure 1 shows the CNOT structure of a logarithmic depth GHZ preparation circuit with an example of a selected pair of qubits together with the coverage. Using a greedy algorithm, we identify the optimal pairs of qubits to connect via a flag to maximize this coverage. This ensures that a small number of flags can effectively monitor the most critical error pathways in the large entangled state.

We note that this flag placement optimization can be viewed as an instance of the classical *maximum coverage problem*: given a collection of sets—each defined by a pair of leaf qubits and the nodes along their respective paths to the least common ancestor in the preparation tree—select k sets whose union is maximized. The objective function $f(S) = |\bigcup_{(i,j)\in S} \mathrm{path}(i,j)|$ is monotone submodular, since the marginal gain from adding a new flag pair diminishes as earlier pairs already cover overlapping portions of the tree. A classical result of Nemhauser, Wolsey, and Fisher [26] guarantees that the greedy algorithm achieves at least a $(1 - 1/e) \approx 0.632$ fraction of the optimal coverage, and this ratio is known to be the best achievable in polynomial time under standard complexity-theoretic assumptions [7]. This provides a formal worst-case performance guarantee for the greedy flag placement strategy we employ.

Finally, we utilize a post-selection scheme: any experimental shot where a flag qubit signals an error (measures '1') is discarded. This yields a "heralded" state with significantly higher fidelity as we will show empirically using noisy simulators and Quantinuum hardware in Sections III B and III C respectively.

B. Estimating fidelity using parity oscillation

The parity oscillation method [13, 23, 32] is a standard approach for estimating state fidelity by decomposing the GHZ state into its population and coherence components. The fidelity F is defined as

$$F = \langle \mathrm{GHZ}|\rho|\mathrm{GHZ}\rangle = \frac{1}{2}(\langle P\rangle + \langle \chi\rangle) \tag{2}$$

where $\langle P\rangle = \langle 0|\rho|0\rangle + \langle 1|\rho|1\rangle$ represents the population, which is estimated directly from computational basis measurements.

The coherence $\langle\chi\rangle = \langle 0|\rho|1\rangle + \langle 1|\rho|0\rangle$ requires measurements in different bases. This is achieved by measuring the global parity operator $\mathcal{P}(\phi)$ after applying a rotation $R_z(-\phi)$ followed by $R_y(-\pi/2)$ to all N qubits

$$\mathcal{P}(\phi) = \langle \bigotimes_{j=1}^{N} (\cos\phi X_j + \sin\phi Y_j)\rangle \tag{3}$$

For an N-qubit GHZ state, this signal oscillates as $C\cos(N\phi + \theta)$. The phase-offset θ is the result of coherent noise acting on the qubits and results in the rotated GHZ state

$$|\mathrm{GHZ}_\theta\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + e^{i\theta}|1\rangle^{\otimes N}). \tag{4}$$

To estimate the coherence C and phase-offset θ, one may perform a non-linear least squares curve fit to the parity data collected at various angles ϕ. Alternatively, a Fourier transform can be applied to extract the signal components

$$I_q = \frac{1}{2(N+1)} \sum_{j=0}^{2N+1} e^{\frac{iqj\pi}{N+1}} \langle \mathcal{P}(\frac{j\pi}{N+1})\rangle \tag{5}$$

In this framework, the absolute coherence is determined by $C = |I_N| + |I_{-N}|$. Both methods are physically equivalent and provide a robust estimate of the state's off-diagonal elements. The state is certified as having genuine multi-partite entanglement if the estimated fidelity

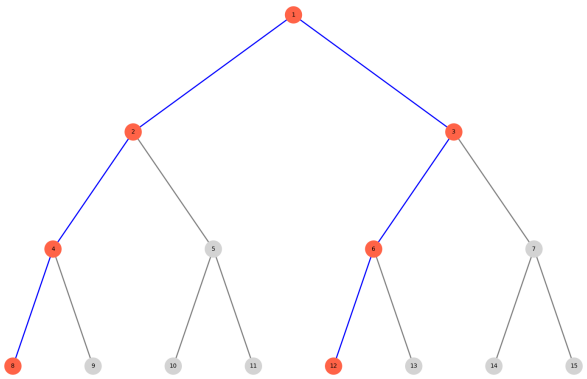

FIG. 1. Binary tree of depth four. The two red leaf nodes are the qubits we use for a parity check, and the paths to their least common ancestor (following the CNOTs) are highlighted in red too. In this case, using just one parity check covers 46.67% of the qubits.

$F > 0.5$. With a non-zero phase-offset θ, the estimated fidelity with respect to the ideal GHZ state will be lower than with respect to the rotated version $|\mathrm{GHZ}_\theta\rangle$. Both have genuine multi-partite entanglement.

To satisfy the Nyquist-Shannon sampling theorem and avoid aliasing, one has to measure the parity signal at $2N$ equally spaced angles $\phi = \frac{k\pi}{N+1}$ for $k = 0, 1, \ldots, 2N-1$.

C. Compressed sensing for sparse signals

Here we follow the reference [9] on compressed sensing, but only focus on the parts necessary for our application.

The compressed sensing problem consists in reconstructing a sparse vector $x \in \mathbb{C}^N$ from

$$y = Ax, \tag{6}$$

where $A \in \mathbb{C}^{M\times N}$ is the so-called measurement matrix and y the measurement values. The vector x is the coefficient vector of the signal f in some orthonormal basis $\{\psi_k\}_k$,

$$f(t) = \sum_k x_k \psi_k(t). \tag{7}$$

Here M is the number of measurement values and we wish to minimize this while still be able to recover x. The efficiency in compressed sensing comes from the randomness of the measurement matrix A. Let $t_1, \ldots, t_M$ be random sampling points and suppose we are given the values

$$y_l = f(t_l) = \sum_k x_k \psi_k(t_l). \tag{8}$$

In this case, the measurement matrix A has entries $A_{l,k} = \psi_k(t_l)$. The task of compressed sensing is to reconstruct f from the samples vector y and to perform this task with as few samples m as possible. We will be mainly interested in the orthonormal basis consisting of $\{\sin(kt), \cos(kt)\}_k$, which is also known as the Discrete Cosine Transform (DCT).

In the context of GHZ verification, the parity oscillation signal $\mathcal{P}(\phi)$ can be expanded as

$$\mathcal{P}(\phi) = C\cos(N\phi+\theta) = C\cos\theta\cos(N\phi) - C\sin\theta\sin(N\phi) \tag{9}$$

We can view this as a DCT of a signal that is extremely sparse, possessing non-zero coefficients only at frequency N. We construct the measurement matrix A of candidate frequencies $k \in \{1, \ldots, N\}$, where N is the system size of the GHZ state. The measurement matrix contains columns for both cosine and sine components

$$A_{k,i} = (\cos(k\phi_i), -\sin(k\phi_i)) \tag{10}$$

We collect data $y_i = \mathcal{P}(\phi_i)$ at M randomly chosen angles ϕ_i, where $M \ll 2N$. The recovery problem is formulated as an L_1-regularized least squares optimization (Lasso)

$$\min_x \frac{1}{2M}||y - Ax||_2^2 + \alpha||x||_1 \tag{11}$$

The L_1 penalty promotes sparsity, forcing most coefficients in x to zero and isolating the true frequency component n_{rec}.

However, empirically, we saw that Lasso estimation introduces a bias (shrinkage) to the coefficients. To correct this, we perform a two-step procedure

1. **Support detection**: Use Lasso to identify the dominant frequency n_{rec} by selecting the index k with the largest coefficient magnitude $\sqrt{a_k^2 + b_k^2}$.

2. **Parameter refinement**: Perform an unregularized Ordinary Least Squares (OLS) fit using only the columns corresponding to n_{rec}. This yields unbiased estimates for $a = C\cos\theta$ and $b = C\sin\theta$.

Finally, the coherence parameters are recovered as $C = \sqrt{a^2 + b^2}$ and $\theta = \arctan(b/a)$.

The theoretical foundation for this efficiency lies in the Restricted Isometry Property (RIP) [4]. For a signal that is s-sparse in the DCT basis of size N, compressed sensing theory guarantees stable recovery with high probability provided the number of random measurements M satisfies:

$$M \geq cs \log(N) \tag{12}$$

where c is a constant. In our case, the signal contains only a single frequency component (plus its conjugate), meaning $s = 2$. Consequently, the required number of measurements scales logarithmically with the system size, $M \propto \log(N)$. This represents a double exponential reduction in measurement overhead compared to full tomography, and an exponential improvement over standard parity oscillation, which typically requires $\mathcal{O}(N)$ samples. This scaling allows us to efficiently verify even very large entangled states.

Crucially, in realistic experimental settings, the parity measurements y_i are corrupted by statistical shot noise and hardware imperfections, yielding $y = Ax + e$ with bounded noise $\|e\|_2 \leq \eta$. The RIP further guarantees that the reconstruction error is bounded by:

$$\|x - x^*\|_2 \leq C_1 s^{-1/2} \|x - x_s\|_1 + C_2 \eta \tag{13}$$

Because the ideal GHZ parity signal is exactly 2-sparse ($s = 2$), the first term (compressibility error) vanishes, and the reconstruction error is dominated entirely by the measurement noise η. Since η scales as $1/\sqrt{S}$ for S experimental shots, this provides a formal guarantee that the compressed sensing estimator will stably converge to the true coherence as the shot count increases, without requiring a dense grid of measurement angles. Furthermore, this noise bound provides theoretical justification for our two-step procedure: the support detection via Lasso is stable under bounded noise η, and once the support is correctly identified, the unregularized OLS fit serves as the optimal unbiased estimator on that support.

We remark that the single-frequency recovery problem we solve has connections to several well-studied areas in classical computer science and signal processing. The sparse Fourier transform literature [14] addresses the general problem of recovering K-sparse frequency representations in sublinear time; for the $K = 1$ case relevant here, classical spectral estimation methods such as Prony's method [33] can recover the frequency algebraically, though they are considerably less robust to noise than the L_1-regularization approach we employ. Our measurement strategy also parallels non-adaptive combinatorial group testing [6], where pooled tests identify a small number of "defective" items in a large population. Each random-angle parity measurement acts as a pooled test that mixes all frequency components, and the logarithmic scaling $M \propto \log(N)$ mirrors the information-theoretic lower bounds for identifying a constant number of active items among N candidates.

III. EXPERIMENTS

To validate our approach, first, we perform numerical simulations to benchmark the accuracy of the compressed sensing estimator and the efficacy of the flag qubit protocol under controlled noise models. Second, we verify our findings on actual quantum hardware, demonstrating high-fidelity GHZ state preparation and efficient verification on Quantinuum's trapped-ion processor.

A. Verification of compressed sensing estimation

Before using the compressed sensing protocol on hardware data, we verify its accuracy and scalability using numerical simulations. The goal is to confirm that the coherence C estimated from a logarithmic number of random samples ($M \propto \log N$) reliably tracks the true state coherence, even in the presence of noise.

1. Methodology

We implemented a verification pipeline comparing the Compressed Sensing (CS)-estimated coherence (C_{est}) against an exact theoretical benchmark (C_{exact}) for GHZ states of size $N \in [5, 40]$. To handle the simulation complexity, we used a hybrid strategy:

- **Small scale ($N \leq 10$)**: We performed full density-matrix simulations using Qiskit's `AerSimulator` under depolarizing noise. C_{exact} was computed directly from the off-diagonal element $|\rho_{0\ldots0,1\ldots1}|$ of the density matrix.
- **Large scale ($N > 10$)**: For larger systems where density matrix simulation is intractable, we utilized a "Fast Emulator". This approach analytically predicts the expected coherence C_{exp} by counting the number of single-qubit (N_{1q}) and two-qubit (N_{cx}) gates in the specific circuit. Assuming a depolarizing noise channel with rates p_{1q} and p_{2q}, the coherence decays as

$$C_{exp} \approx (1 - p_{2q})^{N_{cx}} (1 - p_{1q})^{N_{1q}} \tag{14}$$

  To simulate experimental data, we sample the parity outcomes directly. For a given phase ϕ, the probability of observing even parity is $P_{even}(\phi) = \frac{1}{2}(1 + C_{exp} \cos(N\phi))$. The number of even outcomes in S shots is then drawn from the binomial distribution $n_{even} \sim B(S, P_{even}(\phi))$, yielding a statistical simulation of the measurement process without the cost of vector evolution.

For each N, we performed 100 independent trials. In each trial, the Compressed Sensing (CS) estimator reconstructed the signal frequency and amplitude from $M \approx 5 \ln N$ random phase samples.

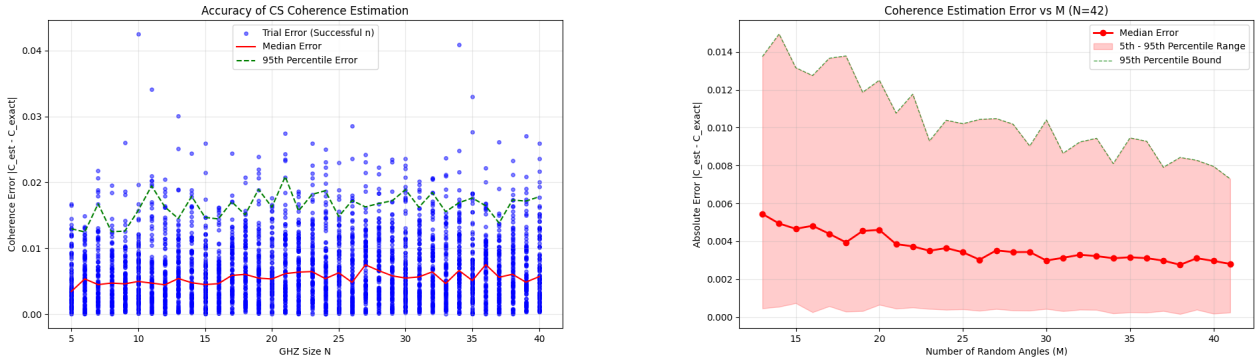

FIG. 2. **Left**: Verification of the compressed sensing estimation. The plot shows the median (red line) absolute error $|C_{est} - C_{exact}|$ and 95th percentile (green dashed line) error, confirming high accuracy across all system sizes. The blue points represent individual trial errors. **Right**: Coherence estimation of a fixed-size GHZ state as a function of the number of chosen random angles. The plot shows the median absolute error $|C_{est} - C_{exact}|$ (red line), the 95th percentile (green dashed line), and the range 5th-95th percentile error.

2. *Results and measurement bias*

The results, summarized in Fig. 2, demonstrate that the CS estimator maintains high accuracy across the entire range of N. The median and 95-percentile error remain bounded, confirming that the logarithmic sampling scaling is sufficient for reliable coherence estimation. The depolarizing error rates that we used here are $p_{1q} = 0$ and $p_{2q} = 0.01$. We explain now why we used perfect single-qubit gates.

The exact coherence measures the state *before* the parity measurement, whereas the CS protocol estimates the visibility *after* applying the necessary rotations ($R_z(\phi)$) and Hadamard gates. These measurement gates introduce additional single-qubit errors spreading across the N qubits, scaling as $(1-p_{1q})^N$. Using noisy single-qubit gates would then add noise to the CS approach, while for the density matrix simulation, that noise would not be present. This was confirmed by running simulations with noisy and perfect single-qubit gates, where the discrepancy appeared with noisy single-qubit gates and disappeared with perfect single-qubit gates. Thus, the CS estimator correctly reports the fidelity, also accounting for the degradation inherent in the measurement process itself.

3. *Frequency identification probability*

The fidelity of the coherence estimation relies on correctly identifying the oscillation frequency n. If the compressed sensing algorithm retrieves the wrong frequency ($n_{rec} \neq N$), the subsequent amplitude fit will characterize noise rather than the signal. Also, we wouldn't be certain that the quantum computer prepared the GHZ state of the correct size: the frequency of the signal is exactly equal to the size of the GHZ state.

We define the "success probability" as the fraction of trials where the recovered frequency matches the true GHZ size $P(n_{rec} = N)$. Fig. 3 illustrates this probability as a function of the number of random samples M for a GHZ state of size $N = 42$. We observe that the success probability rapidly converges as M increases at the start and saturates after M exceeds a heuristic cutoff of $M = 15$ (a bit lower than the $5\log(N)$ threshold used in Figure 2). This empirically confirms the theory that a logarithmic sampling strategy provides sufficient information to uniquely identify the GHZ signal from the sparse spectral domain.

B. Increasing fidelity through parity checks

To show the effectiveness of the error detection protocol before hardware execution, we performed numerical simulations using Qiskit's `AerSimulator`. We modeled a noisy quantum processor with depolarizing error channels, setting the single-qubit gate error rate to $p_{1q} = 0.1\%$ and the two-qubit gate error rate to $p_{2q} = 1.0\%$.

We simulated the preparation and verification of GHZ states for system sizes $N = 10$ under depolarizing and thermal noise. We varied the number of flag qubit pairs k incorporated into the circuit, ranging from $k = 0$ (standard preparation without checks) to $k = 2$ pairs. The "coverage ratio"—defined as the percentage of qubits in the GHZ state whose error paths are monitored by the flags—increases with k.

For each configuration, we estimated the state fidelity using our compressed sensing protocol with $M \approx 5 \ln N$ random measurement angles. The fidelity was calculated as $F = (P+C)/2$, where the population P was measured in the Z-basis, and the coherence C was recovered from the sparse parity oscillation signal. To ensure small error bars, we accumulated $30,000$ shots per measurement

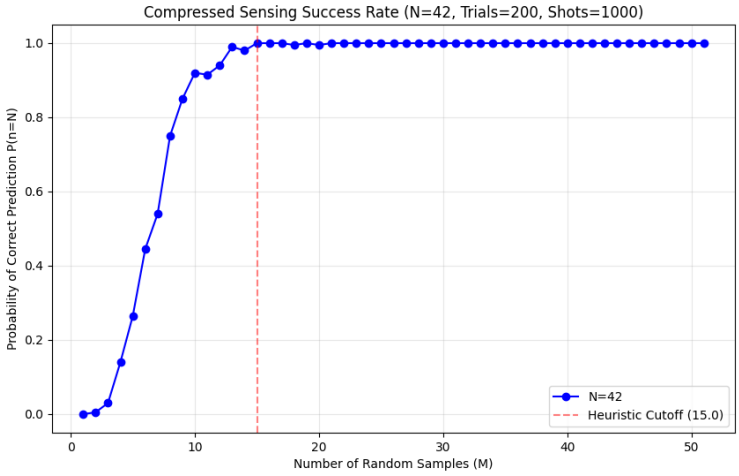

FIG. 3. Success probability of recovering the correct frequency component $n = N$ as a function of random samples M. The vertical line indicates the heuristic threshold 15, beyond which the recovery rate approaches 100%.

circuit.

The results, illustrated in Figs. 4 for the $N = 10$ case, demonstrate a clear advantage to using flag qubits under both depolarizing and (more realistic) thermal noise. As the number of flag pairs increases (corresponding to higher coverage), the post-selected fidelity improves. For example, moving from 0 flags (0% coverage) to 2 flags (90% coverage) yields a marked increase in fidelity, demonstrating that the flags successfully herald the presence of errors. This confirms that maximizing the coverage ratio is a valid proxy for maximizing experimental fidelity.

C. Hardware and emulator results

In this section, we show the results of the experiments performed on the Quantinuum H2-1 quantum computer and on the H2-1E emulator. In Figure 5 the results are shown for a 26 qubit experiments on the H2-1E emulator. In Figure 6 we show the result of a 50-qubit experiment with a varying number of flag qubits on the 56-qubit H2-1 device. In both cases one sees that error detection using flag qubits helps in increased fidelity for the rotated GHZ states.

The divergence between the standard fidelity and rotated fidelity highlights a fundamental trade-off between stochastic and coherent errors. The flag checks effectively detect and filter stochastic bit-flips occurring during state preparation. This post-selection increases the purity and the magnitude of the coherence, thus improving the rotated fidelity. However, executing these extra checks requires additional two-qubit gates and ion transport. In trapped-ion architectures, these extra physical operations introduce coherent over-rotations. For a GHZ state, these local phase errors rapidly accumulate into a global phase offset θ. Because the standard fidelity is penalized by $\cos(\theta)$, this accumulated phase offset suppresses the standard fidelity even as the underlying coherence magnitude improves.

1. *Quantum error mitigation*

While flag qubits detect errors during state preparation, they do not correct for all noise sources present on the hardware. In particular, readout errors systematically bias the observed bit-string distributions, and low-frequency dephasing during idle circuit periods degrades coherence. To address these residual error channels, we layer two quantum error mitigation (QEM) techniques—individually and in combination—on top of the flag qubit protocol: Readout Error Mitigation (REM) [3, 21, 25] and Dynamical Decoupling (DD) [34, 35, 38].

REM corrects measurement bias by applying the inverse of a tensored single-qubit confusion matrix to the post-selected probability vector. Because the confusion matrix factorizes as a tensor product over qubits, the correction scales efficiently to large systems and is applied entirely in classical post-processing—the quantum circuits remain unchanged. DD targets a complementary noise source: during idle time steps where a qubit waits while two-qubit gates act on other qubits, low-frequency dephasing accumulates. Inserting identity-equivalent gate sequences (XX pairs) into these idle windows refocuses the qubit and suppresses decoherence. We use the digital dynamical decoupling implementation from Mitiq [17], applying XX sequences to all idle qubit slots before circuit submission. We evaluate four configurations on 25-qubit GHZ states with $k \in \{0, 2, 4, 6\}$ flag pairs: (i) unmitigated, (ii) REM only, (iii) DD only, and (iv) REM+DD.

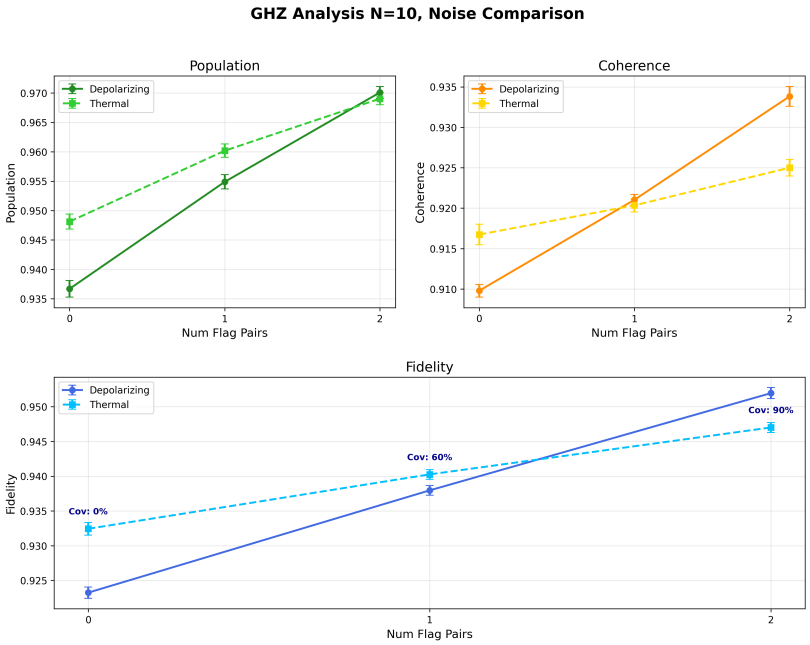

FIG. 4. Simulation results for $N = 10$ GHZ state preparation under depolarizing and thermal noise. The plot shows the estimated fidelity as a function of the number of flag qubits used. Annotation labels indicate the coverage ratio. We observe a clear monotonic increase in population, coherence, and fidelity estimates as more flag qubits are added, confirming that the protocol effectively detects and filters out errors propagating through the circuit.

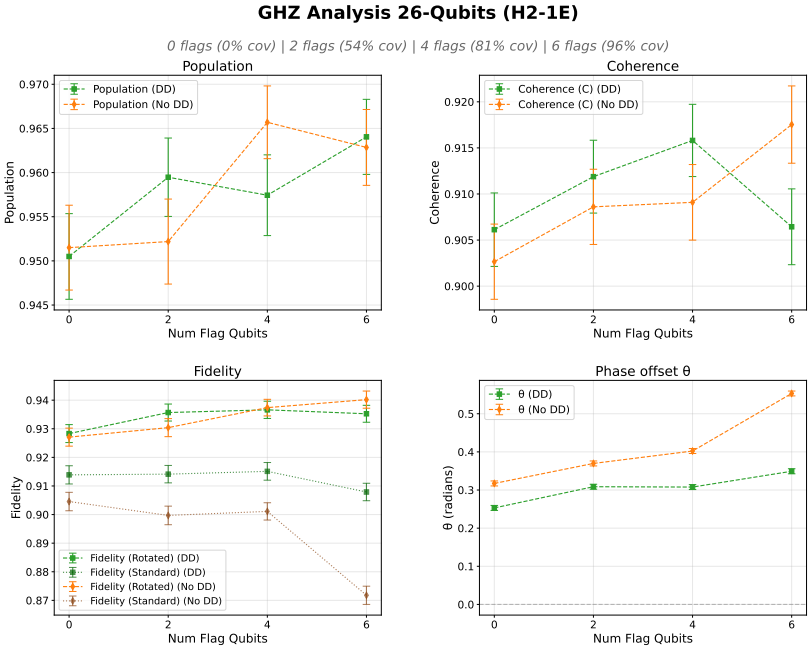

FIG. 5. Simulation results for $N = 26$ GHZ state preparation using Quantinuum H2-1 Emulator with noise. The plot shows the estimated fidelity as a function of the number of flag qubits used. We used 20 random angles at which we sample. The coverage ratio for several flags can be found at the top of the figure. We observe a clear monotonic increase in population and rotated fidelity estimates as more flag qubits are added. However, the standard fidelity is decreasing for which the reason is the increasing phase offset θ, due to unwanted rotations. We have also applied dynamic decoupling (DD).

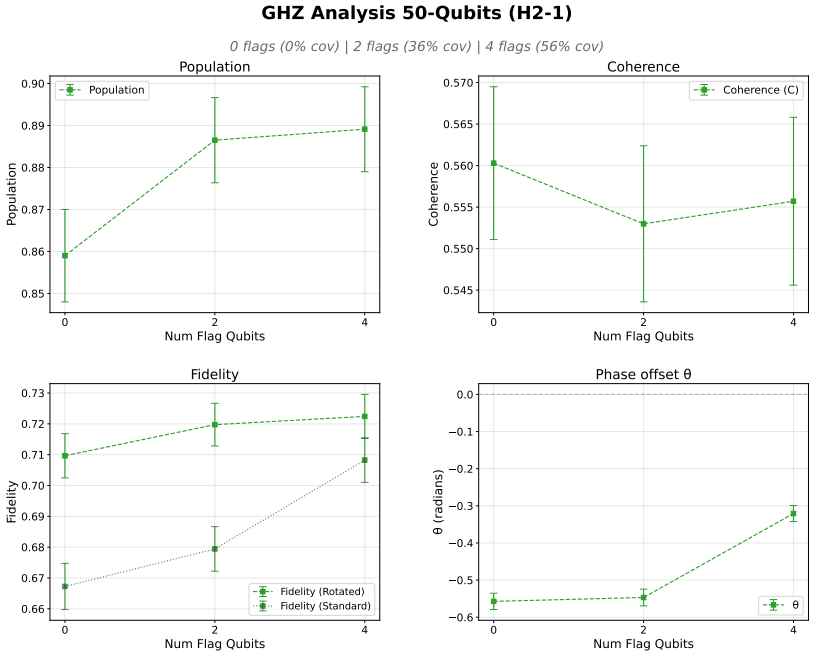

FIG. 6. Estimated fidelity of the 50 qubit standard and rotated GHZ state on the H2 Quantinuum device. We used 1000 shots on each parity oscillation circuit. We used 19 random angles at which we sample. The error bars are quite wide as a result of low number of shots to reduce cost. No error mitigation has been applied here.

As shown in Figure 7, REM provides a massive boost to the estimated fidelity. Because GHZ states are macroscopic, observing the correct population requires all N qubits to be read out perfectly; for $N = 25$, even a 0.2% readout error rate per qubit results in approximately a 5% failure rate per shot. By applying the inverse confusion matrix, REM systematically unbiases this large classical error. In contrast, DD provides only a modest improvement. This is likely because the simplified, memoryless Markovian noise models often used in emulators do not fully capture the correlated, low-frequency drifts that DD is explicitly designed to refocus on physical hardware.

Furthermore, unlike the results on the H2-1 hardware and H2-1E emulator, adding flag qubits on the H2-2E emulator fails to improve the overall fidelity. This indicates that the break-even threshold for error detection is not met under the specific noise model calibrated for H2-2E. If the baseline two-qubit gate error rate is very low, mid-circuit faults are rare. In this regime, the additional gates and transport operations required for the parity checks inject more depolarizing and SPAM errors than they successfully catch, resulting in a net decrease in fidelity.

IV. CONCLUSION

In this work, we have demonstrated a robust and scalable framework for preparing and verifying large GHZ states. By combining flag-based error detection with a compressed sensing verification scheme, we address two critical challenges in quantum computing technology: the sensitivity of large entangled states to noise and the prohibitive cost of full state tomography. Our results confirm that strategically placed flag qubits can enhance state fidelity by heralding errors, while compressed sensing allows us to verify these states with logarithmic measurement overhead. This approach offers a practical pathway for benchmarking GHZ state preparation circuits on next-generation quantum processors.

Our work also reveals connections between quantum state verification and classical algorithmic theory that may guide future improvements. The flag placement optimization is a submodular maximization problem on tree-structured sets, and the binary tree topology of the preparation circuit may admit more efficient exact algorithms—for example, via Euler tour decompositions or tree dynamic programming—that go beyond the $(1 - 1/e)$ approximation guarantee of the greedy approach. On the verification side, the single-frequency structure of the parity oscillation signal connects our method to a broader landscape of spectral estimation techniques and group testing strategies, suggesting alternative recovery algorithms with different noise-robustness tradeoffs. Exploring these classical-quantum algorithmic connections is a promising direction for further improving scalable entanglement verification.

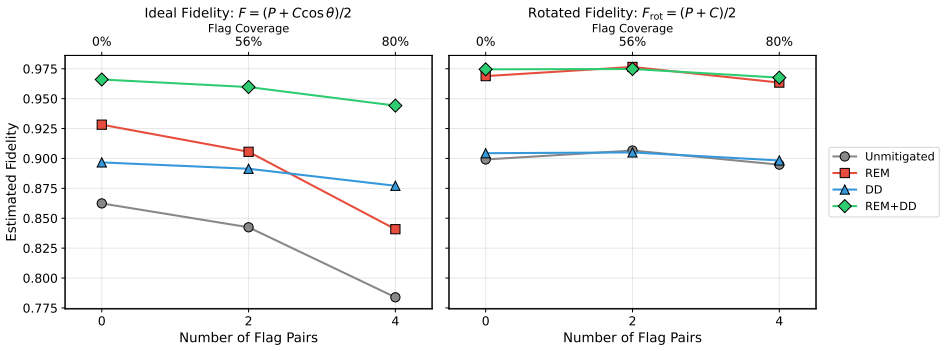

FIG. 7. Comparison of QEM techniques on the Quantinuum H2-2E emulator for 25-qubit GHZ state preparation with compressed sensing verification. The estimated fidelity is shown as a function of the number of flag qubit pairs ($k \in \{0, 2, 4\}$) for four configurations: unmitigated, REM only (analytical correction with $p_{\text{ro}} = 0.2\%$), DD only, and REM+DD. REM provides the largest improvement by correcting measurement bias in both population and coherence estimates. DD shows a modest effect on the emulator, consistent with the emulator's noise model having limited idle-time dephasing. The combination REM+DD yields the highest fidelities across all configurations.

[1] Zehang Bao, Shibo Xu, Zixuan Song, Ke Wang, Liang Xiang, Zitian Zhu, Jiachen Chen, Feitong Jin, Xuhao Zhu, Yu Gao, et al. Creating and controlling global Greenberger-Horne-Zeilinger entanglement on quantum processors. *Nature Communications*, 15(1):8823, 2024.

[2] Joseph Barreto and Daniel Lidar. Compressed sensing shadow tomography. *arXiv preprint arXiv:2602.12518*, 2026.

[3] Sergey Bravyi, Sarah Sheldon, Abhinav Kandala, David C Mckay, and Jay M Gambetta. Mitigating measurement errors in multiqubit experiments. *Physical Review A*, 103(4):042605, 2021.

[4] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.

[5] Marcus P da Silva, Olivier Landon-Cardinal, and David Poulin. Practical characterization of quantum devices without tomography. *Physical Review Letters*, 107(21):210404, 2011.

[6] Ding-Zhu Du and Frank K Hwang. *Combinatorial Group Testing and Its Applications*. World Scientific, Singapore, 2nd edition, 2000.

[7] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[8] Steven T Flammia and Yi-Kai Liu. Direct fidelity estimation from few Pauli measurements. *Physical Review Letters*, 106(23):230501, 2011.

[9] Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Springer New York, New York, NY, 2013.

[10] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004.

[11] Daniel M Greenberger, Michael A Horne, and Anton Zeilinger. Going beyond Bell's theorem. In *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 69–72. Springer, 1989.

[12] David Gross, Yi-Kai Liu, Steven T Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Physical Review Letters*, 105(15):150401, 2010.

[13] Otfried Gühne, Chao-Yang Lu, Wei-Bo Gao, and Jian-Wei Pan. Toolbox for entanglement detection and fidelity estimation. *Physical Review A*, 76(3):030305, 2007.

[14] Haitham Hassanieh, Piotr Indyk, Dina Katabi, and Eric Price. Nearly optimal sparse Fourier transform. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*, pages 563–578. ACM, 2012.

[15] Mark Hillery, Vladimír Bužek, and André Berthiaume. Quantum secret sharing. *Physical Review A*, 59(3):1829, 1999.

[16] Ali Javadi-Abhari, Simon Martiel, Alireza Seif, Maika Takita, and Ken X Wei. Big cats: Entanglement in 120 qubits and beyond. *arXiv preprint arXiv:2510.09520*, 2025.

[17] Ryan LaRose, Andrea Mari, Sarah Kaiser, Peter J Karalekas, Andre A Alves, Piotr Czarnik, Mohamed El Mandouh, Max H Gordon, Yousef Hindy, Aaron Robertson, et al. Mitiq: A software package for error mitigation on noisy quantum computers. *Quantum*, 6:774, 2022.

[18] Dietrich Leibfried, Emanuel Knill, Signe Seidelin, Joe Britton, R Brad Blakestad, John Chiaverini, David B Hume, Wayne M Itano, John D Jost, Christopher Langer, et al. Creation of a six-atom 'Schrödinger cat' state. *Nature*, 438(7068):639–642, 2005.

[19] Haoran Liao, Gavin S Hartnett, Ashish Kakkar, Adrian Tan, Michael Hush, Pranav S Mundada, Michael J Biercuk, and Yuval Baum. Achieving computational gains with quantum error-correction primitives: Generation of long-range entanglement enhanced by error detection. *PRX Quantum*, 6(2):020331, 2025.

[20] Muzhou Ma, Steven T Flammia, John Preskill, and Yu Tong. Learning k-body Hamiltonians via compressed sensing. *arXiv preprint arXiv:2410.18928*, 2024.

[21] Filip B Maciejewski, Zoltán Zimborás, and Michał Oszmaniec. Mitigation of readout noise in near-term quan-

tum devices by classical post-processing based on detector tomography. *Quantum*, 4:257, 2020.

[22] Simon Martiel and Ali Javadi-Abhari. Low-overhead error detection with spacetime codes. *arXiv preprint arXiv:2504.15725*, 2025.

[23] Thomas Monz, Philipp Schindler, Julio T Barreiro, Michael Chwalla, Daniel Nigg, William A Coish, Maximilian Harlander, Wolfgang Hänsel, Markus Hennrich, and Rainer Blatt. 14-qubit entanglement: Creation and coherence. *Physical Review Letters*, 106(13):130506, 2011.

[24] Gary J Mooney, Gregory AL White, Charles D Hill, and Lloyd CL Hollenberg. Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer. *Journal of Physics Communications*, 5(9):095004, 2021.

[25] Paul D Nation, Hwajung Kang, Neereja Sundaresan, and Jay M Gambetta. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum*, 2(4):040326, 2021.

[26] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, 1978.

[27] Ahmed Omran, Harry Levine, Alexander Keesling, Giulia Semeghini, Tout T Wang, Sepehr Ebadi, Hannes Bernien, Alexander S Zibrov, Hannes Pichler, Soonwon Choi, et al. Generation and manipulation of Schrödinger cat states in Rydberg atom arrays. *Science*, 365(6453):570–574, 2019.

[28] Cass A Sackett, David Kielpinski, Brian E King, Christopher Langer, Volker Meyer, Christopher J Myatt, Mary Rowe, Quentin A Turchette, Wayne M Itano, David J Wineland, et al. Experimental entanglement of four particles. *Nature*, 404(6775):256–259, 2000.

[29] Alireza Shabani, Masoud Mohseni, Seth Lloyd, Robert L Kosut, and Herschel Rabitz. Estimation of many-body quantum Hamiltonians via compressive sensing. *Physical Review A*, 84(1):012107, 2011.

[30] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52(4):R2493, 1995.

[31] Chao Song, Kai Xu, Hekang Li, Yu-Ran Zhang, Xu Zhang, Wuxin Liu, Qiujiang Guo, Zhen Wang, Wenhui Ren, Jie Hao, et al. Generation of multicomponent atomic Schrödinger cat states of up to 20 qubits. *Science*, 365(6453):574–577, 2019.

[32] Chao Song, Kai Xu, Wuxin Liu, Chui-ping Yang, Shi-Biao Zheng, Hui Deng, Qiwei Xie, Keqiang Huang, Qiujiang Guo, Libo Zhang, et al. 10-qubit entanglement and parallel logic operations with a superconducting circuit. *Physical Review Letters*, 119(18):180511, 2017.

[33] Petre Stoica and Randolph L Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, Upper Saddle River, NJ, 2005.

[34] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417, 1999.

[35] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A*, 58(4):2733, 1998.

[36] Xi-Lin Wang, Yi-Han Luo, He-Liang Huang, Ming-Cheng Chen, Zu-En Su, Chang Liu, Chao Chen, Wei Li, Yu-Qiang Fang, Xiao Jiang, et al. 18-qubit entanglement with six photons' three degrees of freedom. *Physical Review Letters*, 120(26):260502, 2018.

[37] Ken X Wei, Isaac Lauer, Srikanth Srinivasan, Neereja Sundaresan, Douglas T McClure, David Toyli, David C McKay, Jay M Gambetta, and Sarah Sheldon. Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences. *Physical Review A*, 101(3):032343, 2020.

[38] Jingfu Zhang, Alexandre M Souza, Frederico Dias Brandao, and Dieter Suter. Protected quantum computing: Interleaving gate operations with dynamical decoupling sequences. *Physical Review Letters*, 112(5):050502, 2014.