

Branch-Resolved Characterization of Feed-Forward Error in Dynamic Teleportation via Classical Choi Shadows

Mason Edwards and Prabhat Mishra
University of Florida, Gainesville, Florida, USA

Abstract—Mid-circuit measurement and classical feed-forward are essential primitives for dynamic-circuit teleportation on superconducting quantum processors. However, the error associated with measurement-conditioned corrective operations remains poorly understood when evaluated with respect to individual measurement branches. In this paper, we present a framework for characterizing feed-forward error in dynamic circuit teleportation without losing valuable information related to its behavior across separate branches. We analyze three approaches to applying measurement-conditioned corrections: (i) physical application, (ii) post-processing adjustments, and (iii) a mitigated physical application which utilizes Bit-Flip Averaging (BFA)-based Probabilistic Readout Error Mitigation (PROM). We experimentally reconstruct branch Choi operators via an entangled reference qubit, and validate our physical-application and post-processing Choi-shadow estimators against full tomography of the branch Choi operators. We perform experiments on two physical qubit layouts which differ greatly in mid-circuit measurement readout error, and observe a reversal in the relative order in branch qualities obtained from the post-processing and PROM mitigation strategies. In one physical layout with higher measurement readout error, the operational feed-forward penalty is relatively modest (approximately 0.02-0.03) and PROM produces higher branch qualities than post-processing for every branch. In a separate layout with lower readout error, the operational feed-forward penalty increases to roughly 0.07, and post-processing exceeds PROM for all branch qualities. Our characterization framework can reveal branch-specific error structure and mitigation behavior that state-of-the-art outcome-averaged analyses fail to expose.

Index Terms—Quantum teleportation, mid-circuit measurement, dynamic circuits, quantum characterization

I. INTRODUCTION

The emergence of dynamic circuit features in Noisy Intermediate-Scale Quantum (NISQ) computers has dramatically changed the types of computational processes that can be performed [1]–[5]. For example, recent work has demonstrated that dynamic circuits with mid-circuit measurement and feed-forward can reduce the amount of resources required to perform common operations including the quantum Fourier transform [6], and enable shallow-depth implementations [7]. Before the introduction of mid-circuit measurement and classical feed-forward, conventional quantum circuits only offered terminal measurements, which do not provide support for altering future operations based on their outcome. Many endeavors in quantum computing, including error correction, state preparation, qubit reset, and qubit reuse, depend on the ability to conditionally execute certain quantum operations based on the classical readouts [8]–[11]. However, these efforts face two challenges. Since the measurement process itself can

lead to erroneous readout [12], it can prevent the correct feed-forward operation from being utilized, obscuring the intended post-measurement behavior of the circuit. Moreover, mid-circuit measurements produce both classical and quantum outputs [13]–[15], and effectively branch the system’s evolution into separate outcome-defined branches. This added layer of complexity makes it difficult to characterize error related to measurement-conditioned feed-forward operations.

Quantum teleportation provides a convenient platform to study feed-forward error. In this process, the quantum information of an unknown quantum state is transferred from a sender to a receiver. These two share an entangled resource state, and the sender performs a Bell measurement on the unknown state and their part of the resource. The outcome of this Bell measurement is then classically communicated to the receiver, who performs an outcome-dependent corrective quantum operation to obtain the teleported state [16]. The Bell measurement dynamically evolves the composite quantum system into distinct outcome-defined branches [14], [15], with a separate outcome-dependent corrective operation associated with each branch. Hence, the measurement-induced branches of the teleportation process offer an especially convenient environment for studying quantum dynamics and errors defined by measurement outcomes. There exists a large body of work applying such a paradigm to various areas of quantum information processing, such as quantum error correction, qubit routing and compilation, and quantum networking [17]–[22]. In this paper, we use teleportation as a resource-controlled benchmark to isolate and measure operational feed-forward error with regard to the branching induced by mid-circuit measurements.

Dynamic teleportation has gained attention in NISQ computing [22]. Recent studies have demonstrated a large performance gap between performing corrections actively through dynamic circuits compared to passively in post-selection via measurement categorization [19]. Mid-circuit measurements can be viewed as quantum instruments [15] that produce both classical and quantum output, which can be utilized to determine the subsequent conditional operations [14]. Related work has developed tomography methods specific to measurements, including conditioned process tomography of measurement back-action [23], quantum instrument tomography [14], [15], and quantum non-demolition measurement tomography [24], [25]. The viability of classical shadows for efficiently estimating selected properties of dynamical quantum processes without full process tomography has been demonstrated [26].

Recent work has developed an in situ randomized benchmarking method for measuring mid-circuit measurement-induced error rates in many-qubit circuits [27]. Mitigation of readout error via bit-flip averaging has been shown [28], and PROM has shown the ability to reduce the negative effects of readout error in adaptive circuits with feed-forward [8].

Existing work on teleportation benchmarking and characterization primarily uses outcome-aggregated estimates such as final-state fidelity, average process fidelity, and measures of entanglement preservation [18], [19], which disregard information related to the dynamics specific to individual measurement branches. Prior work has reconstructed process matrices conditioned on individual Bell-measurement outcomes [18], [22]. However, the reported normalized outcome-conditioned process matrices themselves do not encode empirical branch probabilities. Detailed tomographic methods have been developed for comprehensive characterization of mid-circuit measurements [14], [15], [25], yet scalable in situ benchmarking remains primarily focused on aggregate mid-circuit measurement-induced error rates in many-qubit circuits [27].

To the best of our knowledge, the existing literature lacks a branch-resolved, process-level method for characterizing feed-forward error in dynamic teleportation. Such a protocol is necessary to illuminate branching-related discrepancies and provide insight into the extent to which these can be improved by mitigation protocols. In this paper, we address this gap by making the following important contributions:

- 1) We propose a branch-resolved method for characterizing feed-forward error in dynamic quantum teleportation.
- 2) We design a controlled pair of state-preparation quantum circuits that prepare both maximally entangled and non-maximally entangled 4-qubit W_4 states, yet compile to identical circuit depth and gate count while inducing different ideal corrected teleportation channels.
- 3) We demonstrate our framework on IBM Fez, and use two physical qubit layouts to study differing measurement readout conditions. We compare three correction strategies: physical application, post-processing adjustments, and PROM-mitigated physical application.
- 4) We experimentally construct branch Choi operators with classical Choi shadows, validate the physical-correction and post-processing estimators against full tomography of the branch Choi operators, and show that modifying the physical layout reverses the relative branch-quality performance between PROM and post-processing.

II. BACKGROUND: DYNAMIC TELEPORTATION

For teleportation of a single qubit, it is necessary to describe our composite system of qubits in terms of two subsystems, which by convention we refer to as Alice and Bob. Alice's subsystem consists of the input qubit and her part of an entangled resource state which she shares with Bob, whose subsystem also contains the output qubit Alice wishes to teleport the input qubit's state information to. She does this by performing a Bell measurement on her subsystem, which branches the composite system's evolution into one of four

possible subchannels. In this paper, we refer to these as branch maps. Alice's measurement outcome m can be represented by two classical bits via $\Omega = \{00, 01, 10, 11\}$, where $m \in \Omega$. In each measurement-defined branch map $\mathcal{E}_m$, the teleported state can be recovered on Bob's post-measurement output qubit via application of a branch-dependent combination of Pauli operators. Using classical feed-forward, Alice thus sends her observed outcome m to Bob, who can then use his knowledge of m to apply the corresponding sequence of Pauli gates to his qubit and recover the teleported input state [16], in the ideal case where the shared entanglement resource is maximally entangled across the sender-receiver partition [29].

There are two viable methods for applying Pauli corrections to the output state: (i) physically applying the branch-dependent combination of Pauli gates directly to Bob's qubit, and (ii) making adjustments in post-processing. In this paper, we use two types of four-qubit W states as entanglement resources in our teleportation protocol, building on prior work using W states in teleportation [29], [30]. For both resources, the corrective gates for each branch map are:

$$U_m = X^{m_x} Z^{m_z}, \quad (1)$$

where $\mathbb{P} = \{\mathbb{I}, X, Y, Z\}$ is the set of Pauli operators, and $m = m_z m_x \in \Omega$ labels the true branch obtained from Alice's mid-circuit Bell measurement. We define the physical correction channel as:

$$\zeta_m(\rho_{\text{post}}) = U_m \rho_{\text{post}} U_m^\dagger, \quad (2)$$

where ρ_{post} is the uncorrected output state on Bob's qubit following Alice's mid-circuit measurement. This allows us to define the corrected branch map, $\tilde{\mathcal{E}}_m$, as:

$$\tilde{\mathcal{E}}_m = \zeta_m \circ \mathcal{E}_m. \quad (3)$$

Different mathematical tools are needed for describing the aggregate corrected channel $\tilde{\mathcal{E}}$ and the branch maps $\mathcal{E}_m$. While the aggregate channel can be modeled as a completely positive, trace-preserving (CPTP) map, the mid-circuit measurement branches the system's dynamics into a finite set of completely positive (CP), trace-non-increasing (TNI) maps [13]–[15]. Since each branch-specific correction $\zeta_m(\cdot) = U_m \cdot U_m^\dagger$ is a unitary channel, $\tilde{\mathcal{E}}_m$ retains the CP and TNI properties of $\mathcal{E}_m$. Hence, in our branch-resolved characterization of the teleportation process, we cannot rely solely on a single CPTP-channel description and must instead use the quantum instrument representation to describe the branch-resolved dynamics [14], [15]. In this formulation, each outcome m is associated with a CP-TNI map $\mathcal{E}_m$, and for an initial state ρ_{pre} on Alice's input qubit, the probability of obtaining branch m is given by:

$$p(m \mid \rho_{\text{pre}}) = \text{Tr}[\mathcal{E}_m(\rho_{\text{pre}})]. \quad (4)$$

The normalized branch output states for each measurement outcome can then be expressed as:

$$\rho_m = \frac{\mathcal{E}_m(\rho_{\text{pre}})}{p(m \mid \rho_{\text{pre}})}, \quad (5)$$

provided $p(m \mid \rho_{\text{pre}}) > 0$. The mid-circuit measurement is therefore represented by a CPTP instrument map $\mathcal{I}$ from a purely quantum input to a hybrid quantum-classical output [13], [14]:

$$\rho \mapsto \mathcal{I}(\rho) = \sum_{m \in \Omega} \mathcal{E}_m(\rho) \otimes |m\rangle \langle m|, \quad (6)$$

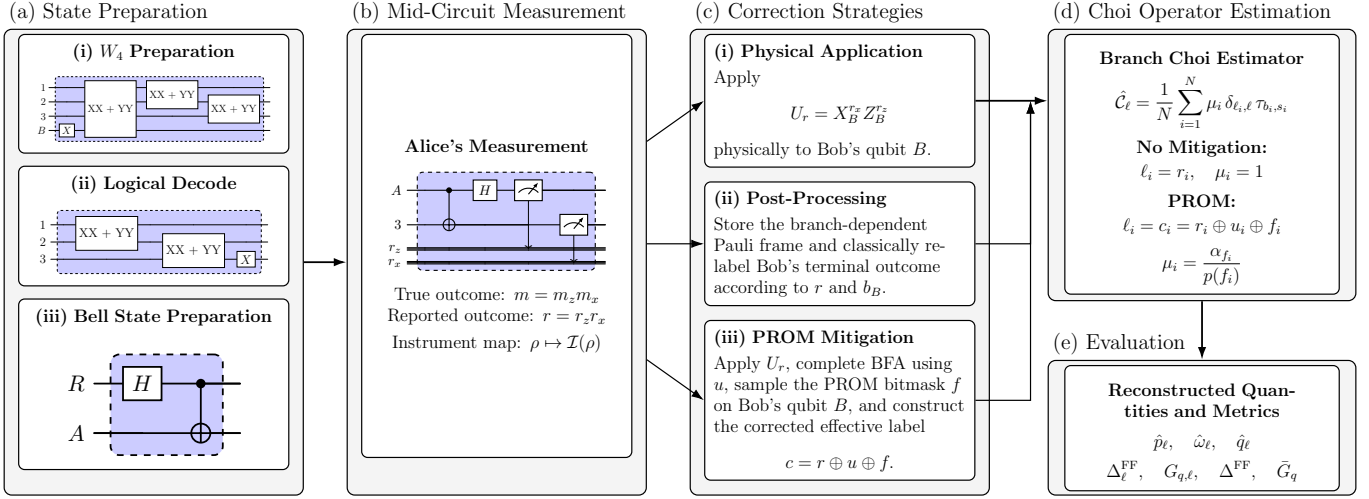


Fig. 1. Branch-Resolved Characterization Framework. **(a)** Sequence of initial state preparations performed preceding the mid-circuit measurement. First, either entangled 4-qubit W_4 resource is prepared on qubits 1, 2, 3, B via the quantum circuit shown in Fig. 2, followed by a decoding of Alice’s resource qubits into their logical basis (Eq. 19) via the quantum circuit shown in Fig. 3. Then, a Bell state is prepared on reference qubit R and input qubit A (Eq. 25). **(b)** Alice’s branch-inducing mid-circuit measurement on qubits A and 3. A preceding bell-basis change is performed, and the measurement outcome readout from qubit A is stored in classical register r_z , while the readout from qubit 3 is stored in classical register r_x . Readout assignment error obscures the ability to determine the true branch $m = m_z m_x$ resulting from the measurement. **(c)** The three methods used to apply the outcome-dependent corrections described in Eq. 1: (i) physical application, (ii) post-processing, and (iii) PROM mitigation. Each method is explained in detail in Section IV. **(d)** Estimation of branch Choi operators. Snapshot operators are constructed following terminal measurements in each shot, which are used in our shadow estimator (Eq. 32) to estimate the observed branch Choi operators for each measurement-induced branch. For the PROM method, a corrected branch label is used, which targets the ideal branches via our mitigated Choi estimator in Eq. 36. **(e)** We reconstruct branch probabilities and branch-quality metrics from these shadow estimators, and use them to estimate feed-forward error and the extent of PROM mitigation as detailed in Section III.

where $|m\rangle \langle m|$ denotes the classical register state storing Alice’s two-bit measurement readout used to apply the appropriate correction. Due to the trace-preserving property of $\mathcal{I}$ [31], for a normalized input state ρ , it follows that:

$$\sum_{m \in \Omega} \text{Tr}[\mathcal{E}_m(\rho)] = 1. \quad (7)$$

Thus, we can represent the aggregate teleportation process in terms of the branch maps via the expression:

$$\tilde{\mathcal{E}} = \sum_{m \in \Omega} \tilde{\mathcal{E}}_m. \quad (8)$$

III. BRANCH-RESOLVED CHARACTERIZATION

In our framework, we define a multipartite teleportation protocol consisting of six total qubits: one reference qubit, labeled R , which we use to construct branch map Choi operators, and five remaining qubits, labeled A , 1, 2, 3, and B , which we use for teleportation. Qubit A contains the input state $|\psi\rangle_A$ we wish to teleport, while qubits 1, 2, 3, and B contain the prepared entanglement resource state. Motivated by prior work demonstrating the utility of multipartite entanglement in teleportation [32], and by later work showing that non-maximally entangled multipartite states can improve controlled teleportation performance [33], we use two types of entangled four-qubit W states to account for effects related to entanglement resource state structure in our teleportation protocol: (i) a primary resource that is maximally-entangled across the Alice-Bob partition, and (ii) a secondary non-maximally entangled control resource state that is not. The general form of a 4-qubit W state on qubits 1, 2, 3, B is:

$$|\Psi\rangle_{123B} = a|1000\rangle + b|0100\rangle + c|0010\rangle + d|0001\rangle. \quad (9)$$

To be physically realizable, any valid W_4 state must satisfy the normalization condition $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. The first state we use is the 4-qubit symmetric W_4 state, which is defined by its equal amplitudes [30] $a = b = c = d = \frac{1}{2}$:

$$|W_4\rangle = \frac{1}{2}(|1000\rangle + |0100\rangle + |0010\rangle + |0001\rangle). \quad (10)$$

This state is properly normalized, and can be used experimentally. However, Bob’s reduced state is not maximally-mixed as $\rho_B = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| \neq \frac{\mathbb{I}}{2}$, so the resource has less than one ebit of entanglement across the Alice and Bob partition and thus does not allow for unit-fidelity teleportation [29]. The ideal teleportation channel resulting from its use thus differs from the identity channel.

The second W_4 resource state is chosen so that Bob’s reduced state is maximally-mixed. Following the correspondence between maximally entangled resource states, singlet fraction, and perfect teleportation fidelity [34], we refer to this resource as the perfect W_4 state. Separate amplitudes can be chosen such that Bob’s reduced state is maximally-mixed [30]. Decomposing Eq. 9 into Alice and Bob’s subsystems yields:

$$|\Psi\rangle_{123B} = (a|100\rangle_{123} + b|010\rangle_{123} + c|001\rangle_{123}) \otimes |0\rangle_B + d|000\rangle_{123} \otimes |1\rangle_B. \quad (11)$$

We can express this in a form resembling a Bell state by defining unnormalized composite states $|\phi_0\rangle_{123} = a|100\rangle_{123} + b|010\rangle_{123} + c|001\rangle_{123}$, and $|\tilde{\phi}_1\rangle_{123} = d|000\rangle_{123}$ for the qubits in Alice’s subsystem. Eq. 11 then becomes:

$$|\Psi\rangle_{123B} = |\tilde{\phi}_0\rangle_{123} \otimes |0\rangle_B + |\tilde{\phi}_1\rangle_{123} \otimes |1\rangle_B. \quad (12)$$

The density matrix for such a state can then be expressed as:

$$\rho = |\Psi\rangle\langle\Psi| = |\tilde{\phi}_0\rangle\langle\tilde{\phi}_0| \otimes |0\rangle\langle 0| + |\tilde{\phi}_0\rangle\langle\tilde{\phi}_1| \otimes |0\rangle\langle 1|$$

$$+|\tilde{\phi}_1\rangle\langle\tilde{\phi}_0|\otimes|1\rangle\langle 0|+|\tilde{\phi}_1\rangle\langle\tilde{\phi}_1|\otimes|1\rangle\langle 1|. \quad (13)$$

For Bob's reduced state to be maximally-mixed across the Alice-Bob partition, the following constraint must be satisfied:

$$\rho_B = \text{Tr}_{123}(\rho) = \frac{\mathbb{I}}{2}. \quad (14)$$

The reduced density matrix for Bob's subsystem can be obtained by taking the partial trace over Alice's qubits in Eq. 13, which provides:

$$\rho_B = (|a|^2 + |b|^2 + |c|^2) |0\rangle\langle 0| + |d|^2 |1\rangle\langle 1|. \quad (15)$$

By defining $p = |a|^2 + |b|^2 + |c|^2$, normalization then requires $1 - p = |d|^2$. We rewrite Eq. 15 as:

$$\rho_B = p |0\rangle\langle 0| + (1 - p) |1\rangle\langle 1|. \quad (16)$$

By setting $p = \frac{1}{2}$, we can satisfy Eq. 14:

$$\rho_B = \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1| = \frac{\mathbb{I}}{2}. \quad (17)$$

Therefore, we must choose amplitudes such that $p = |a|^2 + |b|^2 + |c|^2 = \frac{1}{2}$ and $1 - p = |d|^2 = \frac{1}{2}$. The perfect W_4 resource can thus be defined via the amplitudes $a = b = c = 1/\sqrt{6}$, $d = 1/\sqrt{2}$ [30]:

$$|\tilde{W}_4\rangle = \frac{1}{\sqrt{6}} (|1000\rangle + |0100\rangle + |0010\rangle + \sqrt{3}|0001\rangle) \quad (18)$$

Since $|\tilde{\phi}_0\rangle$ and $|\tilde{\phi}_1\rangle$ are orthogonal, $\langle\tilde{\phi}_0|\tilde{\phi}_0\rangle = p$, $\langle\tilde{\phi}_0|\tilde{\phi}_1\rangle = \langle\tilde{\phi}_1|\tilde{\phi}_0\rangle = 0$, and $\langle\tilde{\phi}_1|\tilde{\phi}_1\rangle = 1 - p$. Thus, we can define normalized logical qubits for Alice's subsystem:

$$|0\rangle_L \equiv |\phi_0\rangle_{123} = \frac{|\tilde{\phi}_0\rangle}{\sqrt{p}}, \quad |1\rangle_L \equiv |\phi_1\rangle_{123} = \frac{|\tilde{\phi}_1\rangle}{\sqrt{1-p}}, \quad (19)$$

with which her four-dimensional logical Hilbert space is $\mathcal{H}_L = \text{span}\{|0\rangle_A \otimes |\phi_0\rangle, |0\rangle_A \otimes |\phi_1\rangle, |1\rangle_A \otimes |\phi_0\rangle, |1\rangle_A \otimes |\phi_1\rangle\}$. Using these and the constraint that $p = \frac{1}{2}$, it can be seen that the perfect W_4 resource is logically equivalent to a Bell state:

$$|\tilde{W}_4\rangle = |\Phi_0^+\rangle_{LB} = \frac{1}{\sqrt{2}} (|0\rangle_L \otimes |0\rangle_B + |1\rangle_L \otimes |1\rangle_B). \quad (20)$$

This form is naturally suited for teleportation, as the two-dimensional logical subspace spanned by $|\phi_0\rangle$ and $|\phi_1\rangle$ can be decoded onto a single logical qubit. These normalized logical basis states are equivalent for both resource states:

$$|\phi_0\rangle_{123} = \frac{1}{\sqrt{3}} (|100\rangle + |010\rangle + |001\rangle), \quad |\phi_1\rangle_{123} = |000\rangle.$$

Inspired by the contributions of Agrawal and Pati [29], we now define a resource-independent orthonormal basis for the four-dimensional logical subspace containing Alice's input qubit and her logical resource qubit:

$$|\eta^\pm\rangle_{A123} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |\phi_0\rangle_{123} \pm |1\rangle_A \otimes |\phi_1\rangle_{123}), \quad (21)$$

$$|\xi^\pm\rangle_{A123} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |\phi_1\rangle_{123} \pm |1\rangle_A \otimes |\phi_0\rangle_{123}). \quad (22)$$

We define her von Neumann projective measurement as $\Pi_{m \in \Omega} = \{\Pi_{00}, \Pi_{01}, \Pi_{10}, \Pi_{11}\}$, where

$$\begin{aligned} \Pi_{00} &= |\eta^+\rangle\langle\eta^+|, & \Pi_{01} &= |\xi^+\rangle\langle\xi^+|, \\ \Pi_{10} &= |\eta^-\rangle\langle\eta^-|, & \Pi_{11} &= |\xi^-\rangle\langle\xi^-|, \end{aligned} \quad (23)$$

TABLE I
MEASUREMENT OUTCOMES AND CORRESPONDING PAULI CORRECTIONS
USED IN THE BRANCH-RESOLVED TELEPORTATION PROTOCOL.

Alice's Measurement	m	Bob's Correction
$ \eta^+\rangle$	00	$\mathbb{I}$
$ \xi^+\rangle$	01	X
$ \eta^-\rangle$	10	Z
$ \xi^-\rangle$	11	XZ

form a complete projective measurement on $\mathcal{H}_L$, as $\sum_{m \in \Omega} \Pi_m = \mathbb{I}_{\mathcal{H}_L}$. Expanding the pre-measurement initial composite state in this basis gives the useful expression:

$$\begin{aligned} |\psi\rangle_A \otimes |\tilde{W}_4\rangle_{123B} &= \frac{1}{2} (|\eta^+\rangle_{A123} \otimes |\psi\rangle_B \\ &+ |\xi^+\rangle_{A123} \otimes X |\psi\rangle_B + |\eta^-\rangle_{A123} \otimes Z |\psi\rangle_B \\ &+ |\xi^-\rangle_{A123} \otimes XZ |\psi\rangle_B). \end{aligned} \quad (24)$$

Alice can thus perform her mid-circuit measurement Π_m and use classical feed-forward to send the outcome $m = m_z m_x \in \Omega$ she observes to Bob, enabling him to apply the corresponding correction U_m listed in Table I and recover the teleported state. The resulting ideal teleportation channel is the identity for the perfect W_4 resource, and for the symmetric W_4 resource is a dephasing channel.

A. Branch-Resolved Choi Operators

To obtain normalized branch Choi states, we maximally entangle a reference qubit, labeled R , with Alice's input qubit A by preparing them in a normalized Bell state:

$$|\Phi_0^+\rangle_{RA} = \frac{1}{\sqrt{2}} (|0\rangle_R \otimes |0\rangle_A + |1\rangle_R \otimes |1\rangle_A) \quad (25)$$

Due to the Choi-Jamiołkowski isomorphism [35], [36], the post-teleportation joint operator on the reference qubit and Bob's output qubit is the branch Choi operator $\mathcal{C}_m$ corresponding to each $\tilde{\mathcal{E}}_m$:

$$\mathcal{C}_m = (\mathbb{I}_R \otimes \tilde{\mathcal{E}}_m) [|\Phi_0^+\rangle\langle\Phi_0^+|], \quad (26)$$

where $m \in \Omega$ is the true branch induced by the mid-circuit measurement. Given that $\tilde{\mathcal{E}}_m$ is both CP and TNI, it follows that $\mathcal{C}_m$ is positive-semidefinite and subnormalized. In the remainder of this paper, we use the term *branch Choi operator* to refer to the subnormalized Choi operator $\mathcal{C}_m$ associated with a single measurement branch. The aggregate corrected Choi operator for the CPTP teleportation process can be expressed as their sum:

$$\mathcal{C} = \sum_{m \in \Omega} \mathcal{C}_m \quad (27)$$

For our entangled-input teleportation process, we define the branch probability p_m as the probability of obtaining the true branch m from Alice's mid-circuit measurement. This can be obtained from taking the trace of the branch Choi operator:

$$p_m = \text{Tr}(\mathcal{C}_m) \quad (28)$$

The normalized branch Choi states for the post-teleportation reference-output subsystem can be obtained via division of each Choi operator $\mathcal{C}_m$ by the respective branch probability:

$$\sigma_m = \frac{\mathcal{C}_m}{p_m} \quad (29)$$

We define the operator $M_m = d[\text{Tr}_B(\mathcal{C}_m)]^T$, which is the branch effect operator associated with the CP-TNI branch map in the Choi representation, and the set $\{M_m\}_{m \in \Omega}$ forms the corresponding Positive Operator-Valued Measure (POVM) [13], [36]. Here, d is the dimensionality of the input state's Hilbert space. For qubit-based experiments, $d = 2$. These M_m allow us to obtain the conditional probability of obtaining a branch m given an input state ρ_A :

$$p(m | \rho_A) = \text{Tr}(M_m \rho_A) \quad (30)$$

Although both the symmetric W_4 and perfect W_4 resource states allow the same logical basis and outcome-dependent corrections, their differing ideal channels cause them to have distinct ideal corrected branch maps and ideal aggregate channels. While we are interested in reconstructing p_m , σ_m , and $\mathcal{C}$ from the true branch Choi operators $\mathcal{C}_m$, the measurement noise obscures our ability to obtain knowledge of the true branch induced by the measurement. It is necessary to instead use a separate label ℓ for the observed branch readouts, whose meaning depends on the correction strategy.

B. Classical Choi Shadows

Completely characterizing dynamical quantum processes is computationally intensive, as the number of measurements required to do so scales exponentially [26], [37], [38] with the number of qubits involved. Our goal experimentally is to estimate the branch-resolved Choi operators $\mathcal{C}_m$ for single-qubit teleportation, and thus the Choi subsystem we aim to reconstruct consists only of the two reference-output qubits R, B . While full tomography of two qubits is relatively tractable computationally, even small increases to the number of qubits would quickly render the task impractical for many experimental scenarios. Alternatively, ShadowQPT [26], [39] has demonstrated that classical-shadow methods can be used to construct stochastic Choi estimators which estimate selected properties of quantum processes using fewer measurements than conventional full process tomography. Thus, we use our two-qubit Choi environment to experimentally validate the accuracy of estimations produced via our non-mitigated Choi shadow estimator (Eq. 32) against full process tomography of the branch Choi operators. For multi-qubit teleportation, and experiments with larger reference-output systems, full tomography of these branch Choi operators would quickly become infeasible, in which case our shadow implementation serves as an insightful reference for efficiently estimating appropriately selected, actionable process information without exhaustive process tomography.

We take a branch-resolved approach to performing classical shadow tomography of the post-teleportation branch Choi operators. For each shadow circuit, a randomized two-qubit Pauli basis $b = (b_R, b_B) \in \{X, Y, Z\}^{\otimes 2}$ for the qubits R, B is generated. Each shot of a given shadow circuit then produces: (1) The classical branch readout obtained from the mid-circuit measurement, and (2) A terminal measurement outcome bitstring $s = (s_R, s_B) \in \{0, 1\}^{\otimes 2}$ for qubits R, B . From these, we construct a two-qubit inverted local Pauli snapshot operator $\tau_{b,s}$ for each shot on qubits R, B [37]:

$$\tau_{b,s} = \bigotimes_{i \in \{R, B\}} (3\Pi_{s_i}^{b_i} - \mathbb{I}_i) \quad (31)$$

Here, $\Pi_{s_i}^{b_i} = U_{b_i}^\dagger |s_i\rangle \langle s_i| U_{b_i}$ is the rank-1 projector associated with the terminal measurement outcome s_i in the Pauli basis b_i , and U_{b_i} is the single-qubit unitary used to perform the measurement in the b_i basis. An important distinction is that for the physical application approach to corrections, the terminal measurement is able to directly measure the teleported output state on B . However, for the post-processing approach, no physical corrective gate is applied following the mid-circuit measurement, and the correction is instead applied through tracked branch-dependent operators in post-processing. Thus, the aggregate estimator built from these snapshots is able to estimate the corrected branch processes, regardless of correction strategy.

We define ℓ_i as the branch label associated with shot i . The shot-specific snapshot operators are then aggregated into branch-specific shadow estimators $\hat{\mathcal{C}}_\ell$ which we use to reconstruct the observed branch Choi operator:

$$\hat{\mathcal{C}}_\ell = \frac{1}{N} \sum_{i=1}^N \mu_i \delta_{\ell_i, \ell} \tau_{b_i, s_i} \quad (32)$$

This estimator groups snapshots by branch via the Kronecker delta $\delta_{\ell_i, \ell}$. The μ_i term is a PROM-specific weight factor we introduce for each snapshot estimator. For our mitigation-free approaches to corrections, we group branches directly by the reported label obtained from mid-circuit measurement, and set μ_i to 1 for each shot. For PROM mitigation, it is determined by the calibrated quasi-probability coefficients and the corresponding PROM sampling distribution.

C. PROM Mitigation and a Mitigated Branch Estimator

While the shadow estimator $\hat{\mathcal{C}}_\ell$ allows the Choi operator to be approximated, unfortunately, readout noise obscures our ability to directly determine the true branch m obtained from each mid-circuit measurement, regardless of correction strategy [8]. For the unmitigated correction strategies, we only have the reported branch label to work with, and cannot target the ideal branches with our estimator. Alternatively, we can utilize Bit-Flip Averaging (BFA) to symmetrize the effective readout channel [28] through specific device calibration circuits and obtain an error syndrome distribution that, combined with PROM quasi-probability re-weighting, allows us to construct an estimator that uses the calibrated symmetrized readout model to target the ideal branch in its approximations.

To calibrate the effective readout error channel for the qubits A and 3 which we perform Alice's mid-circuit measurement on, we first define both a true branch label $t = t_z t_x \in \Omega$ and a bit-flip bitstring $u = u_z u_x \in \Omega$, where the subscripts z and x refer to the Z and X components of the correction to the reported branch label, respectively. We use t as a pre-determined label for the true branch we expect from measurement, and u as a known set of bit flips we apply to qubits A and 3 before measurement with equal probability. If either u_z or u_x is 1, an X gate is applied to qubit A or 3 , respectively. We use a total of 16 circuits to calibrate

over all combinations of $(t, u) \in \Omega \times \Omega$. After the mid-circuit measurement in each circuit, we perform modulo-2 addition (equivalently an XOR operation) between the reported branch label r and u to obtain $\tilde{r} = r \oplus u$. We use this to construct the corresponding readout error syndrome $g = \tilde{r} \oplus t$. From these calibration circuits, we estimate the 4-element symmetrized error-syndrome distribution q_g , which gives the calibrated probability of error syndrome g for the symmetrized mid-circuit measurement on qubits A and 3 :

$$q_g = \frac{1}{|\Omega|^2} \sum_{t, u \in \Omega} p((\tilde{r} \oplus t) = g \mid t, u) \quad (33)$$

The first element q_{00} is the probability of the readout-error-free syndrome $g = 00$, with corresponding basis vector $e_{00} = (1, 0, 0, 0)^T$. Because we have symmetrized these readout errors, the logical operation of the mid-circuit measurement is maintained [8]. The PROM bitmask is defined as $f \in \Omega$, which is applied to qubit B after the feed-forward correction is made. Our syndrome distribution is then used to construct the 4×4 symmetrized readout confusion matrix $Q_{g,f} = q_{g \oplus f}$ [8], which allows us to solve the linear equation $Q\alpha = e_{00}$ and obtain the weight vector $\alpha = (\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11})^T$ of possibly negative PROM quasi-probability coefficients α_f . We use the corresponding L_1 norm $\xi = \|\alpha\|_1 = \sum_{f \in \Omega} |\alpha_f|$ to obtain the non-negative sampling distribution $p(f)$, from which we sample PROM bitmasks [8]:

$$p(f) = \frac{|\alpha_f|}{\xi} \quad (34)$$

In our mitigated approach to corrections, we use our experimentally-calibrated PROM coefficients α_f to define the PROM weight for each shot:

$$\mu_i = \frac{\alpha_{f_i}}{p(f_i)}. \quad (35)$$

After the mid-circuit measurement, we use the reported label r to apply the same correction as in the physical-correction strategy. To complete the BFA Pauli twirl, we apply $X_B^{u_x} Z_B^{u_z}$ to Bob's output qubit, followed by the sampled PROM bitmask f . For shot i , this produces an effective correction label $c_i = r_i \oplus u_i \oplus f_i$, which we use as the branch label ℓ_i for grouping our shadow snapshots. Thus, using the calibrated symmetrized readout model, we form our mitigated branch estimator:

$$\hat{C}_\ell^{\text{PROM}} = \frac{1}{N} \sum_{i=1}^N \frac{\alpha_{f_i}}{p(f_i)} \delta_{\ell_i, \ell} \tau_{b_i, s_i} \quad (36)$$

which approximates mitigated branch Choi operators, indexed by the ideal branch labels.

D. Evaluation Metrics

Having established separate shadow estimators $\hat{C}_\ell$ and $\hat{C}_\ell^{\text{PROM}}$ of the corrected branch Choi operators for our unmitigated and mitigated correction strategies, respectively, the common branch label ℓ thus has a different meaning depending on the correction strategy. For unmitigated approaches, it is the reported branch label r obtained from Alice's mid-circuit measurement. For the mitigated approach, it represents the corrected label $\ell \equiv c = r \oplus u \oplus f$. Thus, we now define the

quantities we reconstruct from these $\hat{C}_\ell$. The estimated branch probabilities are obtained directly from taking the trace:

$$\hat{p}_\ell = \text{Tr}(\hat{C}_\ell) \quad (37)$$

We denote the ideal normalized corrected branch Choi state as ω_ℓ . We define the estimated normalized branch Choi state $\hat{\omega}_\ell$, and the estimated branch quality $\hat{q}_\ell$, respectively, as:

$$\hat{\omega}_\ell = \frac{\hat{C}_\ell}{\hat{p}_\ell}, \quad \hat{q}_\ell = \frac{\text{Tr}(\omega_\ell \hat{C}_\ell)}{\hat{p}_\ell} \quad (38)$$

The branch quality $\hat{q}_\ell$ defines the overlap between our experimentally-reconstructed normalized branch Choi state and the corresponding ideal corrected normalized branch Choi state. We also define the branch-resolved feed-forward penalty:

$$\Delta_\ell^{\text{FF}} = \hat{q}_\ell^{\text{post}} - \hat{q}_\ell^{\text{phys}} \quad (39)$$

which is the defining metric we use to characterize the observed feed-forward error. It is the difference in branch qualities obtained from the physical application approach to corrections and the post-processing approach, where the latter uses feed-forward only classically in post-processing. Positive values indicate that the physical corrective gate has obscured information related to a given measurement-induced branch.

We define the aggregate feed-forward penalty Δ^{FF} as the weighted sum of each branch-specific feed-forward penalty:

$$\Delta^{\text{FF}} = \sum_\ell \hat{p}_\ell^{\text{post}} \Delta_\ell^{\text{FF}} \quad (40)$$

This gives the cumulative branch-quality penalty associated with physical feed-forward operations, weighted by the experimentally-estimated post-processing branch probabilities. Finally, we evaluate the extent of mitigation achieved from PROM within each branch by comparing the PROM branch qualities with those of physical application:

$$G_{q,\ell} = \hat{q}_\ell^{\text{prom}} - \hat{q}_\ell^{\text{phys}} \quad (41)$$

Positive values indicate that PROM improves the reconstructed branch qualities relative to physical correction. In our aggregate metrics plot, we describe the average PROM mitigation across all branch qualities using the following unweighted mean over all four branch labels:

$$\bar{G}_q = \frac{1}{4} \sum_{\ell \in \Omega} G_{q,\ell} \quad (42)$$

We compare the accuracy of our Choi-shadow estimates to full Choi-state tomography using two families of scalar observables. First, we define the estimated branch effect operator:

$$\hat{M}_\ell = d \left[\text{Tr}_B(\hat{C}_\ell) \right]^T \quad (43)$$

where $d = 2$ is the Hilbert space dimension for the input qubit. For each branch $\ell \in \Omega$, our primary set of observables is:

$$\mathcal{O}_\ell^{\text{Primary}} = \{\hat{p}_\ell, \hat{M}_{\ell,X}, \hat{M}_{\ell,Y}, \hat{M}_{\ell,Z}, \hat{q}_\ell\} \quad (44)$$

where $\hat{M}_{\ell,X}$, $\hat{M}_{\ell,Y}$, and $\hat{M}_{\ell,Z}$ are the non-trivial Pauli coefficients of the branch effect operator $\hat{M}_\ell$, given by $\hat{M}_{\ell,P} = \text{Tr}(P \hat{M}_\ell)$. The second set of observables we use is a complete branch Choi observable family, which we define as:

$$\mathcal{O}^{\text{Full}} = \left\{ \text{Tr} \left[(P_R \otimes P_B) \hat{C}_\ell \right] : \ell \in \Omega, P_R, P_B \in \mathbb{P} \right\}. \quad (45)$$

For both observable families, we compare the shadow estimate with the estimate from full tomography using:

$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} (\mathcal{O}_k^{\text{shadow}} - \mathcal{O}_k^{\text{tomography}})^2}, \quad (46)$$

where $\mathcal{K}$ is the set of shared scalar observable labels.

IV. EXPERIMENTS

A. IBM Superconducting Hardware

All experiments were performed on IBM's 156-qubit Heron r2 superconducting quantum processing unit (QPU) IBM Fez. Experiments were designed in Qiskit software, and submitted to the backend as 10-minute batches via Qiskit runtime service. The fixed transpilation seed 98364 was used for reproducibility. All circuits were transpiled with optimization level 1.

B. Logical and Physical Qubit Layouts

Following our convention in Section II, we label the logical qubits in our circuits as $[R, A, 1, 2, 3, B]$, denoting R as the reference qubit, A as the input qubit, 1, 2, 3 as Alice's resource qubits, and B as Bob's qubit. We performed mid-circuit measurements on logical qubits A and 3 in our implemented circuits, and terminal measurements on qubits R and B . The mappings of logical to physical qubits for our layouts are:

Layout 1: $[R, A, 1, 2, 3, B] \rightarrow [7, 6, 3, 4, 5, 16]$

Layout 2: $[R, A, 1, 2, 3, B] \rightarrow [22, 23, 2, 3, 16, 1]$

The mid-circuit measurement pairs were $(A, 3) = (6, 5)$ for layout 1, and $(A, 3) = (23, 16)$ for layout 2. These were chosen specifically so that the physical qubits mapped to A and 3 would have high readout assignment error (RAE) for the first layout, and low RAE in the second. The relevant noise characteristics for each qubit were obtained from the official IBM Fez calibration data available at the time we performed our experiments, and are listed in Table II.

TABLE II
IBM FEZ NOISE CALIBRATION DATA

Physical Qubit	5	6	16	23
Readout Assignment Error	0.05139	0.01831	0.00610	0.00671
T1 Relaxation Time (us)	164.684	265.535	248.926	152.267
T2 Coherence Time (us)	136.006	216.041	165.657	157.805

C. Controlled Resource State Preparation

We first prepare the explicit physical W_4 resource states described in Section II on qubits 1, 2, 3, B with equal circuit depth and gate counts. By applying an X gate to qubit B , and then distributing the resulting amplitude across qubits 1, 2, 3 via the parameterized 2-qubit gates shown in Figure 2, we obtain the explicit form of either the symmetric state given in Eq. 10 or the perfect W_4 state given in Eq. 18, depending on the choice of parameter θ_{B1} . The exact parameters are provided in Table III.

In our teleportation protocol, we perform Alice's measurement as a Bell measurement on qubits A and 3. Thus, after preparing the explicit forms of the resource states on qubits 1, 2, 3, B via the quantum circuit in Fig. 2, we decode the resource state qubits into their logical basis (Eq. 19) so that the

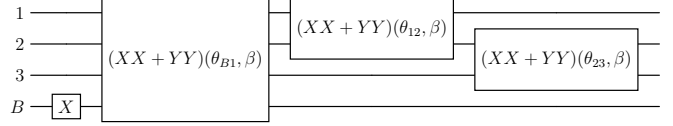


Fig. 2. State preparation quantum circuit. Prepares either entangled W_4 resource state with matched circuit depth and gate count. $\beta = \pi/2$, and $\theta_{B1} = \pi/2$ prepares the perfect W_4 resource, while $\theta_{B1} = 2\pi/3$ prepares the symmetric W_4 resource. θ_{12} and θ_{23} are the same for both resources, as shown in Table III.

two-dimensional logical subspace on qubits 1, 2, 3 is mapped to a single logical qubit on 3. This is accomplished through the decoding circuit shown below in Fig 3, and allows Alice's von Neumann measurement Π_m to be implemented experimentally as a normal Bell measurement on qubits A and 3.

TABLE III
PARAMETERS FOR STATE PREPARATION $XX + YY$ GATES IN FIG. 2.

Entanglement Resource	θ_{B1}	θ_{12}	θ_{23}
Perfect W_4	$\pi/2$	$2 \arccos(1/\sqrt{3})$	$\pi/2$
Symmetric W_4	$2\pi/3$	$2 \arccos(1/\sqrt{3})$	$\pi/2$

The state preparation circuit in Fig. 2 is transpiled together with the logical decoding circuit in Fig. 3, with a combined circuit depth of 36 and size of 74, totaling 36 RZ gates, 28 SX gates, and 10 CZ gates. These depths and counts are the same for both resource states, which allows them to be used as a controlled comparison for analyzing how the differing degrees of entanglement affect the observed error behavior. Once the entangled resource is prepared, we prepare the entangled Bell state in Eq. 25 on qubits R and A . We prepare the logical resource states first in order to minimize the extent of idle decoherence experienced by the teleportation input state preceding Alice's measurement.

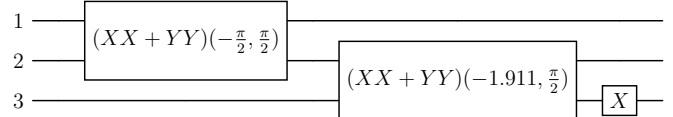


Fig. 3. Decoding quantum circuit. Decodes Alice's resource qubits into a single logical qubit on qubit 3.

D. Teleportation Experiments

For each layout, we perform three sets of experiments corresponding to each of the three correction strategies: (i) physical correction, (ii) post-processing, and (iii) PROM mitigation. Regardless of the strategy, each shadow circuit begins with resource preparation, followed by preparation of a Bell state on qubits R and A . A Bell-basis change is then performed on qubits A and 3, followed by a mid-circuit measurement which produces a reported branch label $r = r_z r_x \in \Omega$. The first bit r_z corresponds to the Z -component of the experimental correction $U_r = X_B^{r_x} Z_B^{r_z}$, and the second bit r_x corresponds to the X -component. For all three strategies, one randomized 2-qubit terminal Pauli measurement basis $b = (b_R, b_B)$ is sampled for each shadow circuit, with each circuit executed for 64 shots. The result of each terminal measurement is labeled

$s = (s_R, s_B)$. The strategy-dependent operations for each shadow circuit are explained below:

a) *Physical Correction*: Immediately following Alice's (Bell) measurement Π_r , the outcome $r = r_z r_x$ is recorded classically and used to apply the corresponding correction U_r physically to qubit B . Afterward, a set of basis-change operations are performed if necessary for measuring R and B in bases b_R and b_B , respectively. The terminal measurement on qubits R, B is performed, and the outcome s is recorded. The local snapshot in Eq. 31 is computed, and the branch label is calculated as $\ell_i = r_i$. Each shot thus contributes $\frac{1}{N} \delta_{\ell_i, \ell} \tau_{b_i, s_i}$ to the non-mitigated estimator (Eq. 32), where N is the total number of shots and $\mu_i = 1$.

b) *Post-Processing*: After Alice's mid-circuit measurement, the outcome r is recorded, however the corresponding correction U_r is not applied physically, and instead is represented as a tracked Pauli frame $F_r = X_B^{r_x} Z_B^{r_z}$ to be used in post-processing. A pre-measurement basis change is performed according to the randomly-chosen bases b_R, b_B , and the terminal measurement is performed. Afterwards, Bob's outcome s_B is flipped depending on (i) the tracked Pauli frame F_r , and (ii) Bob's terminal measurement basis b_B :

- 1) If $b_B = X$, s_B is flipped via classical relabeling in post-processing, only when the tracked Z-component bit is 1.
- 2) If $b_B = Z$, s_B is flipped only when the tracked X-component bit is 1.
- 3) If $b_B = Y$, s_B is flipped only when $r_x \oplus r_z = 1$.

The shot-wise contribution of the snapshot to the non-mitigated estimator is calculated and performed the same as for physical applications.

c) *PROM Mitigation*: For each PROM shadow circuit, in addition to the random terminal measurement basis $b = b_R b_B$, a bit-flip string $u = u_z u_x$ is assigned from a uniformly shuffled list over Ω , and a PROM bitmask $f = f_z f_x$ is sampled from the experimentally calibrated PROM sampling distribution $p(f)$. Before Alice's mid-circuit measurement is performed, the measurement qubits A and 3 are Pauli-twirled via application of $X_A^{u_z} X_3^{u_x}$ in order to symmetrize the readout error channel. The mid-circuit measurement is then performed, and r is recorded and used to physically apply the correction U_r . BFA is then completed via the application of $X_B^{u_x} Z_B^{u_z}$ to Bob's output qubit B before the terminal measurement. The PROM bitmask f is applied to Bob's qubit B , with the terminal measurement performed afterward. For each shot i , a corrected branch label $c_i = r_i \oplus u_i \oplus f_i$ is computed and used as the effective branch label ℓ_i . The snapshot τ_{b_i, s_i} is calculated as in the other strategies using Eq. 31; however, the weight is computed according to Eq. 35. Thus, each shot contributes $\frac{1}{N} \frac{\alpha_{f_i}}{p(f_i)} \delta_{\ell_i, \ell} \tau_{b_i, s_i}$ to the mitigated estimator (Eq. 36).

E. Calibration and Drift Tracking

To account for hardware drift and obtain a meaningful PROM calibration of the error syndrome present in our physical qubit layouts, we partition the experiments into separate batches. For each layout, two batches are performed, with each batch consisting of an initial and a final PROM calibration

of the error syndromes of the qubits used for Alice's mid-circuit measurement. Each calibration uses the complete set of 16 BFA circuits needed to calibrate each combination $(t, u) \in \Omega \times \Omega$ described in Section II, and 2048 shots are executed for each calibration circuit. From these circuits, we obtain the symmetrized readout error syndrome q_g for the qubits $A, 3$, and use it to calculate the PROM quasi-probability coefficients α_f and the experimental bitmask sampling distribution $p(f)$. The α_f used at runtime are determined from the initial calibration, while those used in post-processing are derived from the average of the initial and final calibrations. The same PROM calibration is used for both resources.

F. Shot Budgets

A total of 1,146,880 shots were performed in our main teleportation experiments, which were split into a total of four batches, with two for each physical layout. Within each batch, calibrations were performed both before and immediately after the teleportation experiments. The 16 BFA calibration circuits needed to fully characterize the readout error syndrome were performed for each batch calibration, with 2048 shots per calibration circuit. Thus, 262,144 total shots were used for PROM calibration. Within each physical layout, 147,456 shots were performed for each correction strategy, with 73,728 shots for each resource, as 1152 shadow circuits were performed 64 times each. Thus 442,368 shadow circuit shots were performed per layout, totaling 884,736 shots for the main teleportation experiments. The separate tomography validation experiments on layout 1, performed for both resources and the two unmitigated strategies, contributed an additional 589,824 shots (294,912 for tomography, and 294,912 for shadow circuits).

G. Tomography Validation of Branch Choi Operators

To validate the quality of our shadow estimates, we perform full tomography of the branch Choi operators for the physical-application and post-processing correction strategies. In our tomography baseline, we use all nine two-qubit Pauli measurement bases $\{X, Y, Z\}^{\otimes 2}$ on qubits R and B , with 8192 shots per basis. For each resource and correction strategy, this gives a total of $9 \times 8192 = 73,728$ tomography shots. Counts are grouped by the reported branch label ℓ ; for the post-processing strategy, Bob's terminal measurement outcome is additionally relabeled according to the tracked Pauli frame.

Rather than using Qiskit-native tomography, we reconstruct each normalized branch state on R, B using a fixed-dilution maximum likelihood state tomography approach inspired by prior diluted maximum likelihood tomography [40]. Given measurement projectors Π_j , observed counts n_j , and total counts $N = \sum_j n_j$, we initialize a full-rank density matrix ρ_0 . At iteration k , we compute outcome probabilities $p_j^{(k)} = \text{Tr}(\Pi_j \rho_k)$ and form the state-dependent operator $\mathcal{R}(\rho_k)$:

$$\mathcal{R}(\rho_k) = \frac{1}{N} \sum_j \frac{n_j}{p_j^{(k)}} \Pi_j \quad (47)$$

We update ρ_k as follows:

$$\rho_{k+1} = \frac{G_k \rho_k G_k}{\text{Tr}(G_k \rho_k G_k)}, \quad G_k = \frac{\mathbb{I} + \epsilon \mathcal{R}(\rho_k)}{1 + \epsilon}, \quad (48)$$

using fixed $\epsilon = 1$. Each reconstructed normalized branch density matrix is then multiplied by its empirical branch probability to obtain a branch Choi operator estimate.

We compare these tomography reconstructions with Choi-shadow estimates from eight shadow-circuit budgets: 9, 18, 36, 72, 144, 288, 576, and 1152 executed shadow circuits from the full randomized shadow dataset. Each shadow circuit uses one sampled terminal Pauli basis $b = b_R b_B \in \{X, Y, Z\}^{\otimes 2}$ and is executed with 64 shots, so the total shadow-shot budgets range from 576 to 73,728. For each budget, we compute the RMSE (Eq. 46) between the shadow and tomography estimates over the primary observable family in Eq. 44 and the full branch Choi observable family in Eq. 45. We exclude PROM from this comparison because its mitigated estimator targets ideal branch labels through BFA/PROM quasi-probability weighting, so direct reported-label tomography is not a comparable baseline.

The tomography validation was performed only on layout 1. For each intermediate shadow budget, the RMSE is recomputed over 16 random subsets of executed shadow circuits, and the plotted error bars show the resulting standard deviation. At the full 1152 shadow circuit budget, all shadow circuits are used, so the standard deviation is zero.

H. Statistical Uncertainty Estimation

To estimate shot noise in our sampled shadow circuits, we perform bootstrap resampling from the empirical outcome distribution of each circuit histogram [41], [42]. For each circuit, the sum of the observed counts $\{n_x\}$ is given by $N = \sum_x n_x$, and the empirical probabilities for each outcome are $\hat{p}_x = n_x/N$. For a given shadow circuit with N total shots, classical outcome categories indexed by $x = 1, \dots, K$, and empirical probabilities $\hat{p}_x$, we use NumPy to generate the j^{th} bootstrap replicate count vector $\tilde{\mathbf{n}}^{(j)} = (\tilde{n}_1^{(j)}, \dots, \tilde{n}_K^{(j)})$ as:

$$\tilde{\mathbf{n}}^{(j)} | \mathbf{n} \sim \text{Multinomial}_K(N, \hat{\mathbf{p}}), \quad (49)$$

where $\mathbf{n} = (n_1, \dots, n_K)$ and $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_K)$. This is done independently for each circuit histogram in our shadow dataset. For every bootstrap resample, we recalculate the same branch-resolved estimators (Eq. 32, Eq. 36) and the derived metrics we plot. The PROM coefficients are held fixed during resampling at their analysis-time calibrated values for each block. In our implementation, we use $B = 200$ bootstrap replicates for each shadow dataset, and a 95% confidence level $\gamma = 0.95$ which is used to define upper and lower quantiles via $\alpha = (1 - \gamma)/2$. In our plots, central values are the point estimates from the observed counts, while the error bars indicate the empirical α and $(1 - \alpha)$ quantiles of the bootstrap distribution, corresponding to 95% percentile bootstrap confidence intervals.

V. EXPERIMENTAL RESULTS

Since the reference-output subsystem for our teleportation protocol consists of only two qubits (R, B), full tomography of the branch Choi operators is computationally tractable and provides a conventional baseline for validating our non-mitigated Choi shadow estimator. Our tomography dataset

was acquired from experiments on layout 1 for the physical application and post-processing correction strategies.

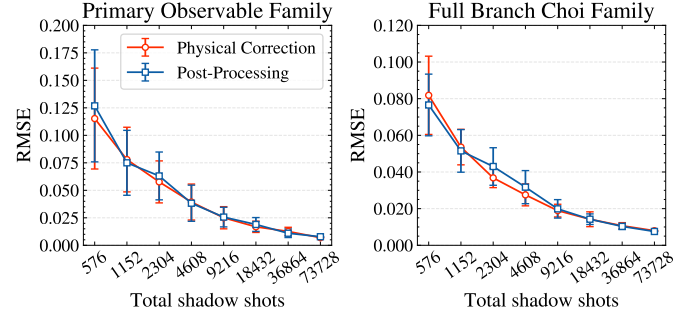


Fig. 4. Shadow estimator validation against full tomography of branch Choi operators on layout 1 using the Perfect W_4 resource. Curves show the RMSE between Choi-shadow estimates and the tomography baseline for the physical correction and post-processing strategies as a function of total shadow shots. Left panel: primary observable family $\mathcal{O}_\ell^{\text{Primary}}$. Right panel: full branch Choi operator family $\mathcal{O}^{\text{Full}}$. Error bars indicate the standard deviation over 16 random subsets of executed shadow circuits at fixed shot budgets.

At the full 73,728 shot budget shown in Fig. 4, the perfect W_4 RMSE values for the primary observable family are 0.00694 (physical application) and 0.00779 (post-processing). For the full branch Choi observable family, the RMSE values for physical application and post-processing are 0.00797 and 0.00754, respectively. The same experiment was performed with the symmetric W_4 state on layout 1, which we do not plot due to page constraints. For the primary observable family, the full-budget RMSE values were 0.01089 (physical application) and 0.01499 (post-processing), and for the full branch Choi observable family, were 0.00838 (physical application) and 0.01108 (post-processing). These results confirm that our non-mitigated Choi-shadow estimator closely matches the full tomography baseline for our two-qubit Choi experiments.

A. Branch-Resolved Feed-Forward Characterization

Having established that our Choi-shadow estimator provides accurate reconstructions of the branch Choi operators, we now present our results from using our branch-resolved estimators to study layout-dependent feed-forward error and PROM mitigation in our main controlled-layout teleportation experiments. The same experimental conditions were used for each layout: the number of shadow circuits, shots per shadow circuit, and calibration methods are identical for experiments in both layouts. Branch quality estimates (Eq. 38) obtained using the perfect W_4 resource state are shown in Fig. 5. In layout 1, PROM mitigation achieves the highest branch qualities for all branches, and physical correction achieves the lowest. Post-processing achieves branch qualities that are worse than PROM but better than physical correction for every branch. In layout 2, however, it is post-processing that achieves the highest branch qualities, exceeding PROM for all branches. Physical correction achieves the lowest branch qualities.

Branch quality estimates obtained using the symmetric W_4 resource state are shown in Fig. 6. In layout 1, the same pattern as Fig. 5 is observed: PROM leads in branch qualities, followed by post-processing, followed by physical correction.

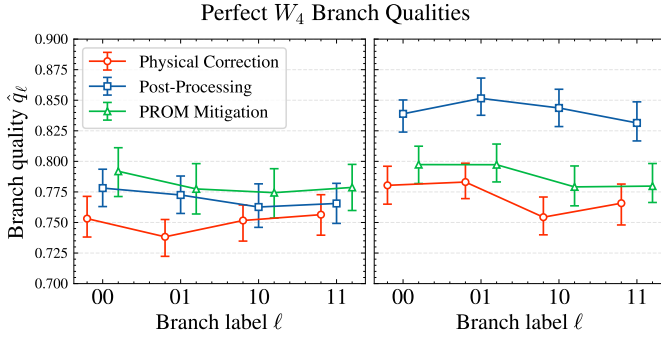


Fig. 5. Estimated branch qualities $\hat{q}_\ell$ obtained experimentally on IBM Fez from each of the three correction strategies with the perfect W_4 resource. Left: results on layout 1. Right: results on layout 2. Error bars represent 95% bootstrap confidence intervals. Relative ordering between post-processing and PROM mitigation effectiveness is reversed between the two layouts.

In layout 2, similar to Fig. 5, the relative effectiveness of PROM compared to post-processing is swapped compared to layout 1. Physical correction has the lowest qualities in every branch, just as with the perfect W_4 resource. Even though the exact branch qualities differ between the two resource states, the observed layout-determined reversal in effectiveness between PROM mitigation and post-processing is preserved.

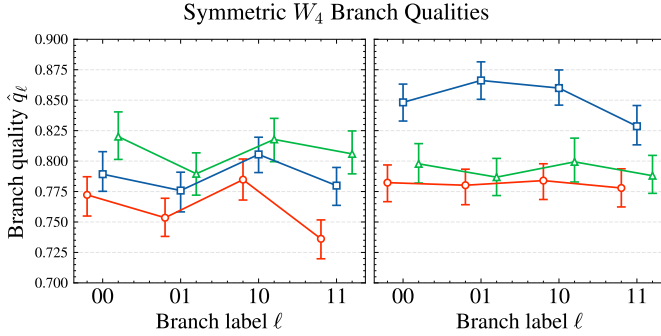


Fig. 6. Estimated branch qualities $\hat{q}_\ell$ for all four branch labels ℓ obtained using the symmetric W_4 resource. Left panel: results from physical qubit layout 1. Right panel: results from physical qubit layout 2. Error bars show the 95% bootstrap confidence intervals. The layout-dependent reversal between PROM mitigation and post-processing is maintained across both resources.

Fig. 7 summarizes our observed aggregated channel metrics. The left panel shows the total feed-forward penalty Δ^{FF} (Eq. 40), and the right panel shows the mean PROM mitigation of branch qualities (Eq. 42), averaged over all four branches.

For layout 1, the total feed-forward penalties are relatively modest at 0.0199 for the perfect W_4 resource, and 0.0259 for the symmetric W_4 resource. For layout 2, the total feed-forward penalties become much larger, at 0.0703 for the perfect W_4 resource, and 0.0698 for the symmetric W_4 resource. The opposite trend is reflected in the mean PROM gain in branch quality $\bar{G}_q$. For layout 1, PROM was able to recover branch qualities much better for each resource state: the perfect W_4 resource resulted in a mean PROM gain of 0.0307, while the symmetric W_4 state yielded 0.0467. Conversely, layout 2 had reduced PROM mitigation, with mean PROM gains of only 0.0175 with the perfect W_4 resource, and 0.0119 for the symmetric W_4 resource.

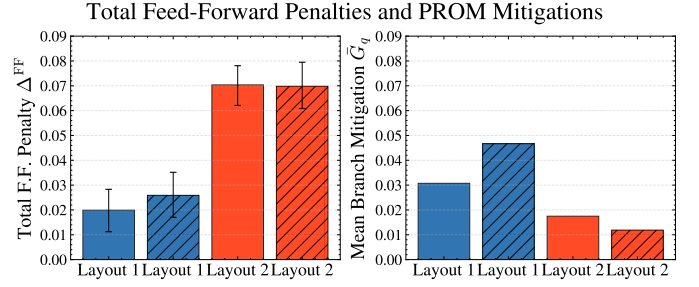


Fig. 7. Total feed-forward (F.F.) penalties Δ^{FF} and mean PROM mitigation of branch qualities $\bar{G}_q$ for each resource state and physical qubit layout. Hatched bars refer to results from symmetric W_4 state experiments, while unhatched bars refer to perfect W_4 results. Error bars in the left panel show the 95% bootstrap confidence intervals. Layout 1 shows lower total feed-forward penalty and noticeably improved PROM mitigation, while layout 2 shows higher total feed-forward penalty and reduced PROM mitigation.

These layout-dependent trends are consistent with our experimentally-obtained PROM calibration data. In our initial and final calibration batches for each layout, the no-error syndrome probability q_{00} is lower in layout 1 than layout 2. For layout 1, the average calibrated readout error syndrome distribution for the two calibration blocks was $(q_{00}, q_{01}, q_{10}, q_{11}) = (0.9308, 0.0509, 0.0174, 0.0009)$, whereas for layout 2 it was $(q_{00}, q_{01}, q_{10}, q_{11}) = (0.9885, 0.0054, 0.0061, 0.0000)$. Thus, higher levels of PROM mitigation were obtained from the layout with the noisier readout channel for the mid-circuit measurement. Conversely, the layout with less readout noise produced a much larger separation between post-processing and physical correction branch qualities, providing a clear view of the overhead associated with physical feed-forward.

VI. CONCLUSION

In this work, we introduced a framework for characterizing feed-forward error in dynamic teleportation via the reconstruction of branch-resolved Choi operators. We also validated the physical-application and post-processing Choi-shadow reconstructions against full tomography of the branch Choi operators on layout 1, providing a conventional tomographic baseline for the estimator used in the main experiments. Using two physical qubit layouts on IBM Fez, three correction strategies, and two entangled W_4 resource states, we found that the level of readout error present in the qubits used for the branch-inducing mid-circuit measurement strongly influences both the cumulative extent of feed-forward error and the relative advantage attained from PROM compared to post-processing. In the layout which has the noisier readout channel, PROM produced the highest branch qualities, and the total feed-forward penalty remained relatively low, while for the layout with the superior readout channel, post-processing produced the highest branch qualities, PROM showed reduced levels of mitigation, and the total feed-forward penalty increased over 2.5 times in magnitude. This pattern persisted across both resource states, demonstrating that our branch-resolved characterization is able to reveal error structure that would be obscured by state-of-the-art outcome-averaged analyses.

REFERENCES

- [1] D. Volya and P. Mishra, “Quantum steering of surface error correcting codes,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2023, pp. 1394–1400.
- [2] D. Volya and P. Mishra, “Quantum benchmarking via random dynamical quantum maps,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2024, pp. 1277–1283.
- [3] D. Volya and P. Mishra, “Quantum spectral clustering of mixed graphs,” in *ACM/IEEE design automation conference (DAC)*, 2021, pp. 463–468.
- [4] S. Sanjaya, H. K. Parvatham, E. Andrews, and P. Mishra, “Controlled steering-based state preparation for adversarial-robust quantum machine learning,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2026.
- [5] S. Sanjaya, D. Volya, and P. Mishra, “Variational quantum algorithms via measurement-induced passive steering,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2024, pp. 481–487.
- [6] E. Bäumer, V. Tripathi, A. Seif, D. Lidar, and D. S. Wang, “Quantum fourier transform using dynamic circuits,” *Physical Review Letters*, vol. 133, no. 15, p. 150602, Oct. 2024.
- [7] C. Cao and J. Eisert, “Measurement-driven quantum advantages in shallow circuits,” *Physical Review Letters*, vol. 136, no. 8, p. 080601, Feb. 2026.
- [8] J. M. Koh, D. E. Koh, and J. Thompson, “Readout error mitigation for mid-circuit measurements and feedforward,” *PRX Quantum*, vol. 7, no. 1, p. 010317, Jan. 2026.
- [9] M. DeCross, E. Chertkov, M. Kohagen, and M. Foss-Feig, “Qubit-reuse compilation with mid-circuit measurement and reset,” *Physical Review X*, vol. 13, no. 4, p. 041057, Dec. 2023.
- [10] D. Volya and P. Mishra, “State preparation on quantum computers via quantum steering,” *IEEE Transactions on Quantum Engineering (TQE)*, vol. 5, pp. 1–14, 2024.
- [11] D. Volya, Z. Pan, and P. Mishra, “Feedback-based steering for quantum state preparation,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2023, pp. 1308–1318.
- [12] Z. Utt, D. Volya, and P. Mishra, “Quantum measurement classification using statistical learning,” *ACM Transactions on Quantum Computing (TQC)*, vol. 5, no. 2, pp. 1–16, 2024.
- [13] E. B. Davies and J. T. Lewis, “An operational approach to quantum probability,” *Communications in Mathematical Physics*, vol. 17, no. 3, p. 239–260, Sep. 1970.
- [14] R. Stricker, D. Vodola, A. Erhard, L. Postler, M. Meth, M. Ringbauer, P. Schindler, R. Blatt, M. Müller, and T. Monz, “Characterizing quantum instruments: From nondemolition measurements to quantum error correction,” *PRX Quantum*, vol. 3, no. 3, p. 030318, Aug. 2022.
- [15] K. Rudinger, G. J. Ribeill, L. C. Govia, M. Ware, E. Nielsen, K. Young, T. A. Ohki, R. Blume-Kohout, and T. Proctor, “Characterizing midcircuit measurements on a superconducting qubit using gate set tomography,” *Physical Review Applied*, vol. 17, no. 1, p. 014014, Jan. 2022.
- [16] 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,” *Physical Review Letters*, vol. 70, no. 13, p. 1895–1899, Mar. 1993.
- [17] C. Ryan-Anderson, N. C. Brown, C. H. Baldwin, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, N. Hewitt, C. Holliman, C. V. Horst, J. Johansen, D. Lucchetti, T. Mengle, M. Matheny, Y. Matsuoka, K. Mayer, M. Mills, S. A. Moses, B. Neyenhuis, J. Pino, P. Siegfried, R. P. Stutz, J. Walker, and D. Hayes, “High-fidelity teleportation of a logical qubit using transversal gates and lattice surgery,” *Science*, vol. 385, no. 6715, p. 1327–1331, 2024.
- [18] M. Baur, A. Fedorov, L. Steffen, S. Filipp, M. P. Da Silva, and A. Wallraff, “Benchmarking a quantum teleportation protocol in superconducting circuits using tomography and an entanglement witness,” *Physical Review Letters*, vol. 108, no. 4, p. 040502, Jan. 2012.
- [19] H. Kang, J. F. Kam, G. J. Mooney, and L. C. Hollenberg, “Teleporting two-qubit entanglement across 19 qubits on a superconducting quantum computer,” *Physical Review Applied*, vol. 23, no. 1, p. 014057, Jan. 2025.
- [20] D. Devulapalli, E. Schoute, A. Bapat, A. M. Childs, and A. V. Gorshkov, “Quantum routing with teleportation,” *Physical Review Research*, vol. 6, no. 3, p. 033313, Sep. 2024.
- [21] H. Rall and M. Tame, “Demonstration of teleportation across a quantum network code,” *Physical Review A*, vol. 110, no. 2, p. 022602, Aug. 2024.
- [22] L. Steffen, Y. Salathe, M. Oppliger, P. Kurpiers, M. Baur, C. Lang, C. Eichler, G. Puebla-Hellmann, A. Fedorov, and A. Wallraff, “Deterministic quantum teleportation with feed-forward in a solid state system,” *Nature*, vol. 500, no. 7462, p. 319–322, Aug. 2013.
- [23] J. Blumoff, K. Chou, C. Shen, M. Reagor, C. Axline, R. Brierley, M. Silveri, C. Wang, B. Vlastakis, S. Nigg, L. Frunzio, M. Devoret, L. Jiang, S. Girvin, and R. Schoelkopf, “Implementing and characterizing precise multiqubit measurements,” *Physical Review X*, vol. 6, no. 3, p. 031041, Sep. 2016.
- [24] L. Pereira, J. García-Ripoll, and T. Ramos, “Complete physical characterization of quantum nondemolition measurements via tomography,” *Physical Review Letters*, vol. 129, no. 1, p. 010402, Jun. 2022.
- [25] L. Pereira, J. J. García-Ripoll, and T. Ramos, “Parallel tomography of quantum non-demolition measurements in multi-qubit devices,” *npj Quantum Information*, vol. 9, no. 1, p. 22, Mar. 2023.
- [26] R. Levy, D. Luo, and B. K. Clark, “Classical shadows for quantum process tomography on near-term quantum computers,” *Physical Review Research*, vol. 6, no. 1, p. 013029, Jan. 2024.
- [27] D. Hothem, J. Hines, C. Baldwin, D. Gresh, R. Blume-Kohout, and T. Proctor, “Measuring error rates of mid-circuit measurements,” *Nature Communications*, vol. 16, no. 1, p. 5761, Jul. 2025.
- [28] A. W. R. Smith, K. E. Khosla, C. N. Self, and M. S. Kim, “Qubit readout error mitigation with bit-flip averaging,” *Science Advances*, vol. 7, no. 47, p. eabi8009, Nov. 2021.
- [29] P. Agrawal and A. Pati, “Perfect teleportation and superdense coding with w states,” *Physical Review A*, vol. 74, no. 6, p. 062320, Dec. 2006.
- [30] B. Pradhan, P. Agrawal, and A. K. Pati, “Teleportation and superdense coding with genuine quadripartite entangled states,” 2007.
- [31] A. Hashim, L. B. Nguyen, N. Goss, B. Marinelli, R. K. Naik, T. Chistolini, J. Hines, J. Marceaux, Y. Kim, P. Gokhale, T. Tomesh, S. Chen, L. Jiang, S. Ferracin, K. Rudinger, T. Proctor, K. C. Young, I. Siddiqi, and R. Blume-Kohout, “Practical introduction to benchmarking and characterization of quantum computers,” *PRX Quantum*, vol. 6, no. 3, p. 030202, Aug. 2025.
- [32] A. Karlsson and M. Bourennane, “Quantum teleportation using three-particle entanglement,” *Physical Review A*, vol. 58, no. 6, p. 4394–4400, Dec. 1998.
- [33] K. G. Paulson and P. K. Panigrahi, “Tripartite non-maximally-entangled mixed states as a resource for optimally controlled quantum teleportation fidelity,” *Physical Review A*, vol. 100, no. 5, p. 052325, Nov. 2019.
- [34] M. Horodecki, P. Horodecki, and R. Horodecki, “General teleportation channel, singlet fraction, and quasidistillation,” *Physical Review A*, vol. 60, no. 3, p. 1888–1898, Sep. 1999.
- [35] M.-D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra and its Applications*, vol. 10, no. 3, p. 285–290, Jun. 1975.
- [36] D. W. Leung, “Choi’s proof as a recipe for quantum process tomography,” *Journal of Mathematical Physics*, vol. 44, no. 2, p. 528–533, Feb. 2003.
- [37] H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” *Nature Physics*, vol. 16, no. 10, p. 1050–1057, Oct. 2020.
- [38] D. Volya, A. Nikitin, and P. Mishra, “Fast quantum characterization via Riemannian optimization,” in *IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2024, pp. 1270–1276.
- [39] J. Kunjummen, M. C. Tran, D. Carney, and J. M. Taylor, “Shadow process tomography of quantum channels,” *Physical Review A*, vol. 107, no. 4, p. 042403, Apr. 2023.
- [40] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, “Diluted maximum-likelihood algorithm for quantum tomography,” *Physical Review A*, vol. 75, no. 4, p. 042108, Apr. 2007.
- [41] B. Efron, “Bootstrap methods: Another look at the jackknife,” *The Annals of Statistics*, vol. 7, no. 1, Jan. 1979.
- [42] B. Efron and R. Tibshirani, “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy,” *Statistical Science*, vol. 1, no. 1, Feb. 1986.