IEEE Open Journal of the Communications Society

Received 30 August 2021; revised 7 October 2021; accepted 15 October 2021. Date of publication 20 October 2021; date of current version 2 November 2021.

Digital Object Identifier 10.1109/OJCOMS.2021.3121183

Quantum Codes in Classical Communication: A Space-Time Block Code From Quantum Error Correction

TRAVIS C. CUVELIER[1,2] (Student Member, IEEE), S. ANDREW LANHAM[2] (Member, IEEE), BRIAN R. LA COUR[2], AND ROBERT W. HEATH, JR.[3] (Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA

[2]Applied Research Laboratories, The University of Texas at Austin, Austin, TX 78713, USA

[3]Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606, USA

CORRESPONDING AUTHOR: T. C. CUVELIER (e-mail: tcuvelier@utexas.edu)

This work was supported by the U.S. Office of Naval Research under Grant N00014-17-1-2107. The work of Travis C. Cuvelier was additionally supported by the University of Texas at Austin Cockrell School of Engineering through an Engineering Doctoral Fellowship.

ABSTRACT We propose a general framework for noncoherent communication using techniques from the field of quantum error correction (QEC). We first propose an approach for analyzing a classical communication channel as a quantum channel, and develop an extension of the QEC conditions to the classical case. We derive a quantum analogue of the noncoherent multiantenna wireless channel. Restricting to the case in which the number of transmitter and receiver antennas are equal and a power of two, we use the framework to develop a family of space-time block codes for noncoherent multiple-input multiple-output (MIMO) communication. Under a Rayleigh fading assumption, we derive the optimal decoder and bound the probability of symbol detection error. We compare our performance to comparable coherent and noncoherent approaches and achieve competitive performance for various antenna geometries and rates.

INDEX TERMS Wireless communications, space-time coding, MIMO, quantum error correction.

I. INTRODUCTION

IN THIS paper we propose a novel signal processing approach to noncoherent communication over linear/affine channels where quantum error correction (QEC) is "emulated" (e.g., using classical hardware or software) in a classical communication system. Generally speaking, noncoherent communication refers to settings where both the transmitter and receiver lack instantaneous channel state information; for example, settings where there is insufficient time and/or resources to form an accurate channel estimate at the receiver. The motivation for our approach stems from the resemblance between quantum communication and classical techniques for noncoherent communication. Like their classical counterparts, quantum communication and information systems are noisy and subject to physical degradation. Techniques from QEC have been developed to mitigate these detrimental effects. Due to fundamental physical limitations, approaches to QEC must operate without knowledge of the underlying system state or the channel realization. In a precise sense, quantum communication is necessarily noncoherent.

We begin by demonstrating an explicit mathematical parallel between quantum error correction and classical noncoherent communication. This leads to a new approach to noncoherent modulation and detection with provable guarantees. In the latter half of this paper, we apply our approach to noncoherent multiple-input multiple-output (MIMO) wireless communication. Taking advantage of our new theoretical results, we use the standard noncoherent MIMO channel model to motivate a particular problem in quantum code design. Solving this quantum code design problem leads to a novel family of noncoherent space-time block codes. Restricting to the case of Rayleigh fading and additive white Gaussian noise (AWGN), we derive the optimal decoder using mathematical techniques from quantum coding theory. We develop a bound on the probability of error and demonstrate via numerical simulations that these codes achieve competitive performance with respect to existing coherent and noncoherent approaches.

A. OUR CONTRIBUTIONS

In this paper, we develop new connections between noncoherent communication and quantum communication. We use these insights to suggest new codes for noncoherent multiantenna communication. Our main contributions are summarized as follows.

1) We propose a novel approach to classical noncoherent communication using mathematical tools from quantum error correction and the theory of quantum error correcting codes. We explicitly derive a mathematical correspondence between classical noncoherent communication over stochastic linear/affine channels and communication over quantum channels. We propose a novel scheme for noncoherent modulation where data is encoded in the mathematical structure of quantum error correcting codes. The approach is sufficiently general, and can be applied to a multitude of classical noncoherent communication problems.
2) We derive theoretical performance guarantees for our approach based on corresponding results in quantum coding theory. These follow from explicit relations between the classical communication problem and problems in quantum code design. As we discuss in the conclusion, these results have additional implications for both quantum and classical information theory.
3) We apply our approach to the noncoherent multiantenna wireless communication setting by developing a family of novel space-time block codes for noncoherent MIMO communication. Restricting to the setting of Rayleigh fading with additive white Gaussian noise, we derive the optimal decoder using the properties of quantum stabilizer codes.
4) We analyze the performance of the derived space-time block code in the Rayleigh fading/AWGN setting. We derive an easy-to-compute bound on the error probability. Finally, we present numerical results indicating that the new family of codes achieves competitive performance for a variety of rates and antenna configurations, with respect to other coherent and noncoherent approaches.

We now proceed with a review of the literature.

B. LITERATURE REVIEW AND CONNECTIONS TO PRIOR WORK

The research in our paper is connected to diverse areas of signal processing and information theory including quantum error correction, quantum-inspired classical signal processing, noncoherent MIMO communication, and noncoherent space-time block coding. In this section, we overview the prior art in these areas, and specifically review the work most relevant to our contributions. In particular, we describe specific relationships between our contributions and those in the literature. We conclude this section with a discussion of the relationship between this paper and our prior work on noncoherent space-time code designs from quantum error correction.

1) QUANTUM ERROR CORRECTION AND QUANTUM-INSPIRED CLASSICAL INFORMATION PROCESSING

In our paper, we propose to approach classical noncoherent communication using techniques derived from quantum error correction, in particular from the theory of quantum error correcting codes. The basic theory of quantum channels, quantum noise, and quantum error correcting codes will be introduced in Section II. At this juncture, it suffices to note that error control protocols in quantum computation and communication systems address novel obstacles compared to classical error correcting codes. Quantum measurement is a dynamical process in that the act of measuring some quantum system causes that system to change [1], [2]. Furthermore, quantum information cannot simply be copied or repeated [1], [2]. In spite of these hurdles, quantum error correcting codes (together with systems to implement encoding and decoding operations) have been developed to enable computation and communication in noisy environments. *Stabilizer codes* are an important class of quantum codes [3]. Stabilizer codes are constructed via group-theoretic notions that are (in a sense) analagous to classical linear block codes [1]. Many families of stabilizer codes find their roots in classical coding theory–the famed CSS (cf. [4]) construction relies on self-dual codes over GF(2). Stabilizer codes based on low-density parity check codes were proposed in [5]. In our work, we look in the other direction and develop a technique for classical error correction using the formalism of quantum stabilizer codes.

Quantum information processing has inspired a wide array of classical signal processing techniques. In [6] and [7], a framework was proposed where quantum measurements were emulated in a classical signal processing environment. Randomized algorithms were developed with applications to various quantization, detection, and estimation problems. In this work, we emulate quantum error correction in classical systems; however, we consider a paradigm that diverges from that in [6] and [7]. We do not consider randomized algorithms, and we do not propose to emulate the quantum measurement process in a randomized way. Instead, we derive (deterministic) optimal detectors using the mathematical toolkit offered by quantum error correction.

A related line of research is *quantum emulation*. Quantum emulation, in our present context, is the use of analog classical signals as a mathematically equivalent representation of a quantum state and the use of analog classical devices to effect unitary transformations and probabilistic measurements upon them. In quantum emulation, the tools of "classical" physics are used to mimic the differential equations governing a quantum system. This is in contrast to quantum simulation, where classical computers are used to perform numerical studies. In more concrete terms, spring-mass systems and passive resistor/inductor/capacitor (RLC) circuits both represent simple harmonic oscillators. Carefully choosing the resistance, capacitance, and inductance of a second order circuit, and viewing a voltage trace on an oscilloscope, can allow one to emulate, in an analog sense, a mechanical spring-mass system. We contrast this to a simulation where (for example) the differential equations governing the mechanical system are discretized and solved using numerical methods. In [8] an approach was proposed where quantum state vectors were represented in the frequency domain. Using ideas from filtering theory and analog circuitry, the so-called quantum emulation device (QED) was shown to emulate a universal quantum computer. A prototype of the system was developed, and, as expected, its performance suffered from the impact of noise. In [9], techniques from quantum error correction were proposed to improve the performance of the QED. In particular, the circuit's additive white Gaussian noise was shown to have an effect analogous to the quantum depolarizing channel. A quantum error correcting code and a corresponding decoding procedure were implemented and shown to improve the fidelity of (emulated) transversal gates by two orders of magnitude. Our paper's theoretical work on the use of quantum codes in classical systems complements the results in [9], and may lead to additional insight into how to deal with noise in future classical systems designed to emulate quantum computation.

2) SPACE-TIME BLOCK CODES FOR NONCOHERENT MULTIANTENNA COMMUNICATION

Space-time codes are designed to exploit the spatial diversity of spreading information over multiple spatial and temporal degrees of freedom. They transmit information over multiple channel realizations and thus provide protection against deep fade events, also known as *outages*. Outages contribute the largest source of symbol decoding errors in nominally slowly-varying, rich-scattering environments like Rayleigh fading [10]. There is a fundamental tradeoff between *diversity*, the rate at which the outage probability decreases at a high signal-to-noise ratio (SNR), and *multiplexing*, a proxy for data rate [11], [12]. Space-time coding approaches can combine spatial multiplexing and diversity to exploit different points on the diversity-multiplexing tradeoff, as explicitly discussed in [13]. While space-time coding was an active research area in the early 2000s, interest declined in light of results suggesting that for sufficiently long packet sizes and for sufficiently high packet error tolerances, modern mobile broadband systems should dedicate all available degrees of freedom to multiplexing [10], [14]. In recent years, however, this paradigm has been reexamined in light of modern developments in short blocklength information theory (cf. [15]) and rapid deployment of emerging technologies that will rely on ultra-reliable, low-latency communication (URLLC) [16]. Recent analyses have suggested directing the degrees of freedom offered by multiple antennas towards reducing block error probability, rather than towards gaining spectral efficiency [10], [17]. In particular, [17] demonstrated that full-diversity space-time codes, used as an inner code and combined with powerful outer codes across different channel realizations, can enable communication near theoretical limits in URLLC-like settings.

In the short-blocklength setting, obtaining channel state information may incur a non-negligible overhead. To address this, space-time codes have been developed for *noncoherent communication*, where neither the receiver nor the transmitter are assumed to form instantaneous channel state information. The ergodic, capacity-achieving signaling structure for high-SNR noncoherent communication in Rayleigh fading is Unitary Space-Time Block Coding (USTBC) [18]. In USTBC, codebooks carry an interpretation as packings on the Grassmann manifold [19]. This insight has led to a variety of noncoherent codebook designs based on the construction of Grassmannian packings. Codebooks for noncoherent space-time codes have been generated using frame theory [20], numerical optimization techniques [21], and orthogonal designs [22]. Other noncoherent design methods can be viewed as procedures for transforming coherent codebook designs into noncoherent ones [23], [24]. Our present work straddles the algebraic and optimization approaches to codebook design. Our quantum-inspired approach to space-time coding uses the algebraic structure of quantum error correcting codes to map numerically-optimized Grassmann packings in lower-dimensional spaces to those in higher-dimensional spaces (cf. Section IV-B). This property leads to intuitive, easy-to-compute bounds on our code's pairwise error probability. Notably, the difficulty of computing the bounds do not depend on the number of antennas. In particular, for low rates, the quality of our family of codes can be easily assessed for arbitrary numbers of antennas.

The algebraic space-time code construction in [25] is most relevant to our present work. In [25], Grassmann packings were developed using Reed-Mueller codes as well as formalism from the Pauli group–a matrix group that plays a key role in the theory of quantum stabilizer coding. The packings in [25] are constructed algebraically and optimized for spectral efficiency and error-rate at a moderate signal-to-noise ratio. The codes presented in our present work are distinct from [25] in both construction and viewpoint; in our present work, we cast noncoherent communication as a quantum channel and solve a resulting code design problem. The construction of the space-time block code is not purely algebraic. Furthermore, our construction is optimized for high SNR performance, while the Reed-Mueller codes are better suited for lower SNRs.

Other noncoherent code designs, such as differential space-time codes (DSTCs), depart from a capacity-motivated design approach. DSTCs encode information in the relationships between pairs of successive codeword transmissions. The codes depend only on the coherence length of the channel, requiring no other statistical characterization. Fast-fading environments, therefore, may not support effective communication with DSTCs. Generally, higher-performance designs are achievable for longer channel coherence lengths, as longer sequences of transmissions can encode more information; though such DSTCs are increasingly vulnerable to carrier frequency offset, which can gradually degrade the signal quality due to phase mismatch. Early designs in the differential coding framework were based on transformations of USTBCs [26], [27]. In contrast, many proposed designs leverage algebraic properties of matrix groups to create new codebooks [28], [29], [30], [31]. A separate family of DSTC designs extends the theory of orthogonal designs to propose novel codebooks [32], [33]. Differential space-time coding has more recently received renewed attention in the context of URLLC [34]. In contrast to other approaches, including the space-time codes derived in this manuscript, differential approaches can be argued to implicitly perform channel estimation due to their use of a known "reference matrix" at the beginning of transmission. This may be undesirable from a physical-layer security perspective.

In this paper, we focus on performance in the regime of extremely short block lengths. We consider the uncoded bit error rates to be the relevant figures of merit; e.g., we do not consider coding across multiple channel realizations/coherence times. Several constructions, including [18], [28], and [35] derive explicit bounds on the symbol error rate (SER) via Chernoff bounds of pairwise error probabilities (PEP) to explicitly characterize performance. In [36] the maximum likelihood (ML) receivers are defined for both coherent and non-coherent communication (including differential modulation and unitary signaling) in the setting where the MIMO channel is a zero-mean Gaussian random matrix. A suboptimal generalized likelihood ratio test is proposed for detection in environments with correlated fading. The test is equivalent to the ML rule in Rayleigh fading. Under mild assumptions on the channel correlation matrix, the suboptimal receiver performs perfectly at infinite SNR. We use an approach similar to these works in Section III-C, where we employ ideas from QEC to derive an analogous detection rule guaranteed to perform perfectly at infinite SNR for a rich class of channel distributions. In Section VI, we demonstrate that this rule corresponds with the ML rule in Rayleigh fading. We demonstrate that, by the nature of our space-time code design, the decoder derived in this paper has reduced computational complexity (for equivalent rates and numbers of antennas) with respect to the decoder proposed in [36].

3) CONNECTION TO OUR PRIOR WORK

This work is based, in part, on our previous work [37]. In our prior work we used ideas from quantum error correction, in particular the mathematical formalism of quantum stabilizer coding, to design a noncoherent space-time block code specialized to the case of two antennas at both the transmitter and the receiver. We derived the optimal decoder assuming Rayleigh fading and AWGN for this restricted case. In the present manuscript, we develop a quantum-inspired approach for noncoherent communication over linear channels. In contrast, our past work only considered communication over the standard narrowband block fading model from MIMO wireless communication. We also develop new theoretical guarantees for this general approach, none of which appeared in [37]. As in [37], we consider applications of quantum error correction to the design of noncoherent space-time block codes. In this present work, however, we generalize the 2×2 space-time block code from [37] to a family of codes for $2^k \times 2^k$ systems, where k is a positive integer. We extend the optimal decoder from [37] to the general case. The performance evaluation in [37] was limited to a numerical study. In the present work, we derive a bound on pairwise error probability that applies to the entire family of codes and present new numerical results (with larger numbers of antennas) for the generalized code construction.

C. ORGANIZATION AND NOTATION

In Section II, we review necessary background material on QEC, quantum channels, and stabilizer codes. In Section III we propose to apply techniques from QEC to classical noncoherent communication settings over linear channels. We propose to view the channel's input as a quantum state, and, given the channel model, define a corresponding *emulated* quantum channel model that is amenable to analysis and code design via the techniques from QEC. In Section IV we apply the formalism developed in Section III to the problem of MIMO wireless communication and develop a novel space-time block code (STBC). The mathematics of QEC lead to a simplified derivation of the maximum likelihood (ML) decoder for the STBC, as detailed in Section VI. In Section VII, we proceed with analysis and numerical results which demonstrate that the family of proposed space-time codes achieve full diversity and competitive performance (e.g., superior for some rates and antenna configurations) with respect to other coherent and noncoherent schemes.

We use non-bold letters to denote scalars. We use bold lowercase Latin letters $\mathbf{a}$ to denote column vectors, and use bold uppercase Latin letters $\mathbf{A}$ or Greek letters $\mathbf{\Lambda}$, $\mathbf{\Gamma}$ to denote matrices. A positive-semidefinite matrix is a symmetric matrix without any negative eigenvalues. We abbreviate positive-semidefinite by PSD and write $\mathbf{A} \succeq \mathbf{B}$ if $\mathbf{A} - \mathbf{B}$ is PSD. We denote the set of $n \times n$ complex PSD matrices $\mathbb{S}_{+}^{n}$ We denote the element in the i^{th} row and k^{th} column of a matrix $\mathbf{A}$ by $[\mathbf{A}]_{i,k}$. In general we denote the $k \times k$ identity matrix by $\mathbf{I}_k$. We use $\mathbf{0}$ to denote the zero matrix or vector (which should be clear from context). We use

$\mathrm{Tr}(\mathbf{A})$ to denote the trace, $\det(\mathbf{A})$ the determinant, $\mathbf{A}^{\mathrm{T}}$ the transpose, and $\mathbf{A}^{\mathrm{H}}$ the conjugate transpose. The commutator between two square matrices of the same size $\mathbf{A}$ and $\mathbf{B}$ is denoted $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. We say that $\mathbf{A}$ and $\mathbf{B}$ commute if $[\mathbf{A}, \mathbf{B}] = \mathbf{0}$ and anti-commute if $\mathbf{AB}+\mathbf{BA} = \mathbf{0}$. We use $|a|$ to denote the absolute value of a scalar, the cardinality of a set, or the order of an algebraic group. We abbreviate "independent and identically distributed" by IID. We use $\mathbb{E}[\,\cdot\,]$ to denote expectation. We use $\otimes$ to denote the tensor product when acting on vector spaces (i.e., $\mathbb{C}^2 \otimes \mathbb{C}^2$) and to denote the Kronecker product when acting on vectors or matrices. We use $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to denote a complex circularly symmetric normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. For a predicate P, we use $\mathbb{1}_P$ as the indicator function onto the set where P is true. We use $\mathbf{A} \overset{\text{a.s}}{=} \mathbf{B}$ to denote equality almost surely (i.e., $\mathbb{P}[\mathbf{A} = \mathbf{B}] = 1$). If $\mathbf{A} = c\mathbf{B}$ where $c > 0$, we say $\mathbf{A}$ is (directly) proportional to $\mathbf{B}$ and we write $\mathbf{A} \propto \mathbf{B}$. If for some positive stabilizer group generators, $\mathbf{A} \overset{\text{a.s.}}{=} c\mathbf{B}$, we write $\mathbf{A} \overset{\text{a.s}}{\propto} \mathbf{B}$.

II. QUANTUM INFORMATION AND QUANTUM STABILIZER CODES

In an effort to make this paper relatively self-contained, we begin by reviewing some background material from quantum information processing. First, we introduce *qubits*, or quantum bits, which are the natural generalization of bits in quantum computing. We describe the Pauli group, an important matrix group in quantum information processing, which we later use to design a space-time code. We continue with a description of *quantum operations* which characterize how systems of qubits can evolve. We conclude with a discussion of QEC, in particular, the powerful class of stabilizer codes. For more detailed descriptions, the reader is directed to [1].

A. QUANTUM STATES, UNITARY EVOLUTION, AND MEASUREMENTS

A quantum state is represented by a complex unit vector $\boldsymbol{\psi}$. The vector encodes a family of probability mass functions (PMFs) that describes the outcomes of all possible experiments, or observations, that can be performed on the system of interest. It turns out that these probabilities do not depend on the *global phase* of the vector. In this sense, a quantum state is represented by the equivalence class of vectors $\{\mathbf{x} \mid \mathbf{x}{=}e^{j\theta}\boldsymbol{\psi}\ \forall\ \theta \in [0, 2\pi)\}$. A state is equivalently represented by a complex PSD matrix $\boldsymbol{\Lambda} = \boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{H}}$ with $\mathrm{Tr}(\boldsymbol{\Lambda}) = 1$. The matrix $\boldsymbol{\Lambda}$ is known as a *density matrix* (or density operator). Density matrices can also be used to describe a quantum system whose state is random. Assume that the state of the system is the random unit vector $\boldsymbol{p}$. The density matrix describing the system is

$$\boldsymbol{\Lambda} = \mathbb{E}\big[\boldsymbol{p}\boldsymbol{p}^{\mathrm{H}}\big]. \tag{1}$$

Like before, $\boldsymbol{\Lambda}$ fully encodes the PMFs describing the outcomes of all possible experiments that can be performed on the system. When the density matrix $\boldsymbol{\Lambda}$ describes a system whose state vector is deterministic, $\mathrm{rank}(\boldsymbol{\Lambda}) = 1$. In this case, the state is said to be *pure*. If $\mathrm{rank}(\boldsymbol{\Lambda}) > 1$, the state is said to be *mixed* and the underlying state vector is random. Any matrix $\boldsymbol{\Lambda} \succeq 0$ with $\mathrm{Tr}(\boldsymbol{\Lambda}) = 1$ is a "valid" density matrix. It follows from diagonalization that there exists a subset of pure states (the normalized eigenevenctors of $\boldsymbol{\Lambda}$) and a probability measure over them (the corresponding eigenvalues), such that $\boldsymbol{\Lambda}$ is equivalent to an expectation over pure state density matrices (e.g., (1)). This decomposition is not necessarily unique. In fact, there are generally many different (both discrete and continuous) distributions over pure states that give rise to identically distributed measurement outcomes and thus equivalent mixed states.

A qubit encodes the state of a two-level quantum system, such as the spin of an electron or the polarization of a photon. In this case $\boldsymbol{\psi} \in \mathbb{C}^2$ and $\boldsymbol{\Lambda} \in \mathbb{C}^{2\times 2}$. The state space of a system of several qubits is the tensor product of the state spaces of the individual systems; the state space of an n-qubit system is given by $\mathbb{C}^{2^n} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. Consider a quantum system composed of n single-qubit subsystems, each in some state $\boldsymbol{\psi}_i \in \mathbb{C}^2$. The state vector of the whole system is the Kronecker product of the state vectors of the individual subsystems via $\boldsymbol{\psi} = \boldsymbol{\psi}_1 \otimes \boldsymbol{\psi}_2 \otimes \cdots \otimes \boldsymbol{\psi}_n$. Analogously, the system's density matrix is given by $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_1 \otimes \boldsymbol{\Lambda}_2 \otimes \cdots \otimes \boldsymbol{\Lambda}_n$. It turns out that the space of state vectors for an n-qubit system includes all unit vectors $\boldsymbol{\psi} \in \mathbb{C}^{2^n}$, not just the set of density matrices that can be written as Kronecker products of smaller systems (this follows from the linearity of quantum mechanics). Analogously the set of n-qubit density matrices is the set of PSD matrices $\boldsymbol{\Lambda} \in \mathbb{C}^{2^n \times 2^n}$ with $\mathrm{Tr}(\boldsymbol{\Lambda}) = 1$.

Closed quantum systems evolve according to unitary transformations [1]. If the density matrix at time 0 is $\boldsymbol{\Lambda}_0$, the density matrix at time t is given by $\boldsymbol{\Lambda}_t = \mathbf{U}_t \boldsymbol{\Lambda}_0 \mathbf{U}_t^{\mathrm{H}}$ for some unitary matrix $\mathbf{U}_t$. In quantum computing, one frequently encounters the concept of applying a unitary gate to a quantum system (analogous to how Boolean operations are applied in classical computing). This turns out to be possible to good approximation [1].

One of the fundamental properties of quantum systems is that the act of observing, or measuring, a quantum system causes it to change [1]. Furthermore, the density operator of a quantum system cannot be directly observed. Instead, measurements provide only incomplete information about the quantum state. One important class of quantum measurements is *projective measurements*. On an n-qubit system, a projective measurement is described by r mutually orthogonal projection matrices $\{\boldsymbol{\Pi}_0, \boldsymbol{\Pi}_2, \ldots \boldsymbol{\Pi}_{r-1}\}$, where $1 \le r \le 2^n$ and

$$\sum_{i=0}^{r-1} \boldsymbol{\Pi}_i = \mathbf{I}_{2^n}. \tag{2}$$

The outcome of a projective measurement is a random integer X, where X takes values in $\{1, 2, \ldots, r\}$. The probability of observing an outcome of i when the state of the system is $\mathbf{\Lambda}$ is given by the Born rule via [1]

$$\mathbb{P}[X = i] = \mathrm{Tr}(\mathbf{\Pi}_i \mathbf{\Lambda}). \tag{3}$$

Given that the outcome i occurs, after measurement, the state's density matrix is given by $\hat{\mathbf{\Lambda}}$ where

$$\hat{\mathbf{\Lambda}} = \mathbf{\Pi}_i \mathbf{\Lambda} \mathbf{\Pi}_i / \mathbb{P}[X = i]. \tag{4}$$

Again, we emphasize that one of the most salient features of quantum measurement is that the measurement probabilities do not depend on the global phase(s) of the state. No experiment can distinguish between two quantum systems with equivalent density matrices, and thus the global phase is not "measurable" in any physical sense. In effect, if two pure states are equivalent up to a global phase, they carry the same information. In other words, the information carried by a pure state depends only on the subspace spanned by its state vector. In Section III, we will propose to encode classical information in the mathematical structure of a constellation of pure quantum states. A key requirement of such constellations will be that no two elements can span the same subspace.

B. PAULI OPERATORS AND THE PAULI GROUP

The Pauli group is an important matrix group that is often used in quantum information science [1], [2]. The Pauli matrices are given by $\mathbf{I}_2$ as well as

$$\mathbf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5}$$

These matrices are both Hermitian and unitary. Taken together, they span $\mathbb{C}^{2\times 2}$. The Pauli matrices generate the so-called 1-qubit *Pauli group*, $G_{\mathcal{P}_1}$, under matrix multiplication. The Pauli group contains all of the Pauli matrices, including the multiplicative factors ±1 and $\pm j$ [1].

The general n-qubit Pauli group, $\mathcal{G}_{\mathcal{P}_n}$ consists of matrices in $\mathbb{C}^{2^n\times 2^n}$ that are generated by the set of n-fold Kronecker products of the Pauli matrices [1]. An element of $\mathbf{G} \in \mathcal{G}_{\mathcal{P}_n}$ has the form

$$\mathbf{G} = z\mathbf{R}_0 \otimes \mathbf{R}_1 \otimes \ldots \otimes \mathbf{R}_{n-2} \otimes \mathbf{R}_{n-1}, \tag{6}$$

where $z \in \{\pm 1, \pm j\}$ and $\mathbf{R}_i \in \{\mathbf{I}_2, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ for $i \in \{0, \ldots, n-1\}$. The elements of the Pauli group $\mathcal{G}_{\mathcal{P}_n}$ are sometimes called n-qubit Pauli matrices. We will frequently consider the subset of the Pauli Group composed of the 4^n elements of $\mathcal{G}_{\mathcal{P}_n}$ with a leading coefficient (e.g., z in (6)) of 1. We denote this set $\mathcal{P}_n$. This set is a Hermitian, unitary, orthogonal (with respect to the trace inner product) basis for $\mathbb{C}^{2^n\times 2^n}$. Further, if $\mathbf{W} \in \mathcal{P}_n$, then $\mathrm{Tr}(\mathbf{W}^{\mathrm{H}}\mathbf{W}) = \mathrm{Tr}(\mathbf{I}_{2^n}) = 2^n$.

The Pauli matrices are unitary and are either Hermitian (with real eigenvalues) or (owing to the multiplicative factors of $\pm j$) skew-Hermitian (with imaginary eigenvalues). They are thus orthogonally diagonalizable with eigenvalues of $\{\pm 1, \pm j\}$. We will use the fact that operators in $\mathcal{G}_{\mathcal{P}_n}$ either commute or anti-commute with one another. The following identities can be verified by direct substitution

$$\begin{aligned} \mathbf{XY} &= -\mathbf{YX} && (7a)\\ &= j\mathbf{Z} && (7b)\\ \mathbf{XZ} &= -\mathbf{ZX} && (7c)\\ &= -j\mathbf{Y} && (7d)\\ \mathbf{YZ} &= -\mathbf{ZY} && (7e)\\ &= j\mathbf{X}. && (7f) \end{aligned}$$

Therefore, each of $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ anti-commutes with the other two.

The identities in (7), along with the bilinearity of the Kronecker product, can be used to derive conditions for two matrices in $\mathcal{G}_{\mathcal{P}_n}$ to either commute or anti-commute. Let $\mathbf{A}, \mathbf{B} \in \mathcal{G}_{\mathcal{P}_n}$ with $\mathbf{A} = a\mathbf{A}_0 \otimes \mathbf{A}_1 \otimes \ldots \otimes \mathbf{A}_{n-1}$ and $\mathbf{B} = b\mathbf{B}_0 \otimes \mathbf{B}_1 \otimes \ldots \otimes \mathbf{B}_{n-1}$, where $\mathbf{A}_i, \mathbf{B}_i \in \{\mathbf{I}_2, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ and $a, b \in \{\pm 1, \pm j\}$. Let $\mathbf{A}_i\mathbf{B}_i = \mathbf{C}_i$, and let $q_i = 0$ if $\mathbf{A}_i$ and $\mathbf{B}_i$ commute and $q_i = 1$ if they anti-commute. Let $s = \sum_{i=0}^{n-1} q_i$. It turns out that whether or not $\mathbf{A}$ and $\mathbf{B}$ commute with one another depends only on s. To see this, note that by definition the products $\mathbf{AB}$ and $\mathbf{BA}$ can be written as

$$\begin{aligned} \mathbf{AB} &= (ab)\mathbf{A}_0\mathbf{B}_0 \otimes \mathbf{A}_1\mathbf{B}_1 \otimes \ldots \otimes \mathbf{A}_{n-1}\mathbf{B}_{n-1} && (8a)\\ &= (ab)\mathbf{C}_0 \otimes \mathbf{C}_1 \otimes \ldots \otimes \mathbf{C}_{n-1}, && (8b) \end{aligned}$$

and

$$\begin{aligned} \mathbf{BA} &= (ab)\big((-1)^{q_0}\mathbf{C}_0\big) \otimes \ldots \otimes \big((-1)^{q_{n-1}}\mathbf{C}_{n-1}\big) && (8c)\\ &= (-1)^s\mathbf{AB}. && (8d) \end{aligned}$$

Thus $\mathbf{BA} = \mathbf{AB}$ if s is even, and $\mathbf{BA} = -\mathbf{AB}$ otherwise. Since all (square) matrices commute with themselves and with the identity matrix, there is a convenient shorthand to determine if $\mathbf{A}$ and $\mathbf{B}$ commute or anti-commute. One simply needs to count the indices i where both $\mathbf{A}_i \neq \mathbf{I}_2$ and $\mathbf{B}_i \neq \mathbf{I}_2$ and also $\mathbf{A}_i \neq \mathbf{B}_i$ via

$$\begin{aligned} s &= \sum_{i=0}^{n-1} \mathbb{1}_{\mathbf{A}_i\neq\mathbf{I}_2, \mathbf{B}_i\neq\mathbf{I}_2, \mathbf{A}_i\neq\mathbf{B}_i} && (9)\\ &= \sum_{i=0}^{n-1} \mathbb{1}_{[\mathbf{A}_i,\mathbf{B}_i]\neq 0}. && (10) \end{aligned}$$

If s is even, $\mathbf{A}$ and $\mathbf{B}$ commute, otherwise, they anti-commute.

C. QUANTUM CHANNELS: MODELS FOR QUANTUM NOISE

Quantum channels can be used to describe the dynamics of open quantum systems [1], [2]. While not completely general, they can be used to describe the quantum extension of the discrete memoryless channel [1]. It turns out that all physical quantum channels from n to n qubits can be written as a deterministic map $\mathcal{E}$ from the space of n qubit density matrices to the space of n qubit density matrices with a very

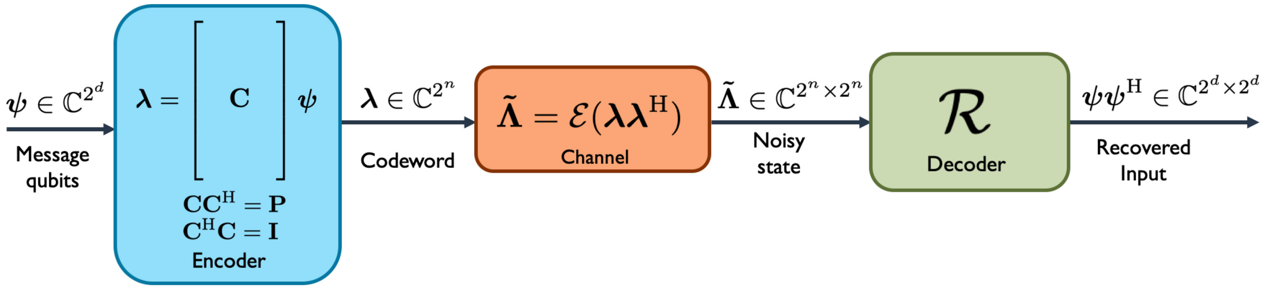
$\psi \in \mathbb{C}^{2^d}$
Message qubits
$\boldsymbol{\lambda} = \left[\mathbf{C}\right] \psi$
$\mathbf{C}\mathbf{C}^{\mathrm{H}} = \mathbf{P}$
$\mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}$
Encoder
$\boldsymbol{\lambda} \in \mathbb{C}^{2^n}$
Codeword
$\tilde{\boldsymbol{\Lambda}} = \mathcal{E}(\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{H}})$
Channel
$\tilde{\boldsymbol{\Lambda}} \in \mathbb{C}^{2^n \times 2^n}$
Noisy state
$\mathcal{R}$
Decoder
$\psi\psi^{\mathrm{H}} \in \mathbb{C}^{2^d \times 2^d}$
Recovered Input

FIGURE 1. The "system model" of QEC. An exploded view of the decoder is depicted in Fig. 2.

particular form [2]. For some set of matrices $\mathbf{E}_i \in \mathbb{C}^{2^n \times 2^n}$, a quantum channel $\mathcal{E}$ can always be written

$$\mathcal{E}(\boldsymbol{\Lambda}) = \sum_i \mathbf{E}_i \boldsymbol{\Lambda} \mathbf{E}_i^{\mathrm{H}}, \tag{11a}$$

where

$$\sum_i \mathbf{E}_i^{\mathrm{H}} \mathbf{E}_i = \mathbf{I}_{2^n}. \tag{11b}$$

The matrices $\mathbf{E}_i$ are known as either *Kraus operators* or *error operators* [1].

Counterintuitively, $\mathcal{E}$, a deterministic map between density matrices, can be used to describe a wide variety of stochastic channels. In quantum systems, both the "inputs" and "outputs" of a channel are considered to be random state vectors (equivalently, density matrices). Since the only way to obtain any information from a quantum state is to measure it, it is sufficient to describe how a channel transforms the input measurement distributions to those at the output. For example, consider a channel that applies a random unitary matrix $\mathbf{U}$ to the input $\boldsymbol{\Lambda}$ according to the probability kernel $f_{\mathbf{U}|\boldsymbol{\Lambda}}$. If the realization of the input Λ is unknown, the density matrix of the output is given by $\boldsymbol{\Lambda}_+ = \mathbb{E}[\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}|\boldsymbol{\Lambda}] = \int_{\mathbf{U}} \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{H}} f_{\mathbf{U}|\boldsymbol{\Lambda}}(\mathbf{V}|\boldsymbol{\Lambda}) d\mathbf{V}$. It turns out that the map from $\boldsymbol{\Lambda}$ to $\boldsymbol{\Lambda}_+$ can be written in the form (11) for some finite set of operators $\mathbf{E}_i$. Again, since many probability distributions over pure states give rise to the same mixed state (cf. Section II-A), it follows that different stochastic kernels can give rise to equivalent quantum channels. Any channel with error operators $\mathbf{E}_i$ is equivalent to a channel that maps an input state $\boldsymbol{\Lambda}$ to $\mathbf{E}_i \boldsymbol{\Lambda} \mathbf{E}_i^{\mathrm{H}}$ with probability $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}} \mathbf{E}_i \boldsymbol{\Lambda})$.

A map $\mathcal{E}$ has an error operator decomposition like (11a) if and only if it is both *convex linear* and *completely positive* [1]. Given that $\mathcal{E}$ can be written in the form (11a), (11b) holds if and only if $\mathcal{E}$ is *trace preserving* [1]. The latter of these is the most intuitive. Assuming that $\mathcal{E}(\boldsymbol{\Lambda})$ is a density matrix for all input density matrices $\boldsymbol{\Lambda}$, the requirement (11b) ensures that $\mathrm{Tr}(\mathcal{E}(\boldsymbol{\Lambda})) = 1$. The former requirements (1) and (2) are similar, although physically motivated: for a quantum channel, (2) ensures that if the input state is the pure state $\boldsymbol{\Lambda}_i$ with probability p_i, the output density matrix is given by $\sum_i p_i \mathcal{E}(\boldsymbol{\Lambda}_i)$ (equivalent to the density matrix describing the situation where $\mathcal{E}(\boldsymbol{\Lambda}_i)$ is the output with probability p_i) [1]. The former condition (1) ensures that if $\boldsymbol{\Lambda} \succeq \mathbf{0}$ then $\mathcal{E}(\boldsymbol{\Lambda}) \succeq \mathbf{0}$, and further ensures that this property holds if $\mathcal{E}$ acts on a subsystem of a larger quantum system [1].

D. QUANTUM ERROR CORRECTION AND STABILIZER CODES

Quantum error correcting codes have a great deal in common with classical linear block codes. A system-level depiction of QEC is shown in Fig. 1. Like classical linear block codes, a quantum code is defined by a subspace C. The code subspace is parameterized by a projection matrix. It will be our convention to denote a projector onto a quantum code by $\mathbf{P}_0$. A code that encodes d qubits in n qubits is a 2^d dimensional subspace of the 2^n dimensional n-qubit state space. In other words, the projector onto the code has $\mathbf{P}_0 \in \mathbb{C}^{2^n \times 2^n}$ and $\mathrm{rank}(\mathbf{P}_0) = 2^d$. Note that the code subspace is exactly isomorphic to the state space to d qubit pure states. We refer to n as the "blocklength" in qubits, and d as the number of "message" or "logical" qubits. An n-qubit pure state $\boldsymbol{\lambda}$ is "in the code" parameterized by $\mathbf{P}_0$ if $\mathbf{P}_0 \boldsymbol{\psi} = \boldsymbol{\lambda}$. Likewise, a density matrix $\boldsymbol{\Lambda}$ is in the code if its row/column spaces are in the range of $\mathbf{P}_0$ (e.g., $\boldsymbol{\Lambda} = \mathbf{P}_0 \boldsymbol{\Lambda} \mathbf{P}_0$).

The physical process by which a lower-dimensional quantum system is encoded in a higher-dimensional system is nontrivial and beyond our scope. For our purposes, it suffices to consider encoding pure states. Let $\mathbf{C}$ be a matrix whose columns are an orthonormal basis for the subspace spanned by $\mathbf{P}_0$, i.e., let $\mathbf{C} \in \mathbb{C}^{2^n \times 2^d}$ be any matrix such that

$$\mathbf{C}\mathbf{C}^{\mathrm{H}} = \mathbf{P}_0 \text{ and } \mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}_{2^d}. \tag{12}$$

We call $\mathbf{C}$ a generator matrix for the code. If $\boldsymbol{\psi} \in \mathbb{C}^{2^d}$ is the d-qubit message to be encoded, the resulting codeword is

$$\boldsymbol{\lambda} = \mathbf{C}\boldsymbol{\psi} \tag{13}$$

where $\boldsymbol{\lambda} \in \mathbb{C}^{2^n}$.

The codeword passes through a noisy channel $\mathcal{E}$ parameterized by some set of error operators $E = \{\mathbf{E}_i\}$ (cf. Section II-C). The outcome of the noisy channel is (generally) a mixed state. After passing through the channel, the receiver can perform measurements on the received noisy state and apply unitary transformations conditioned on the

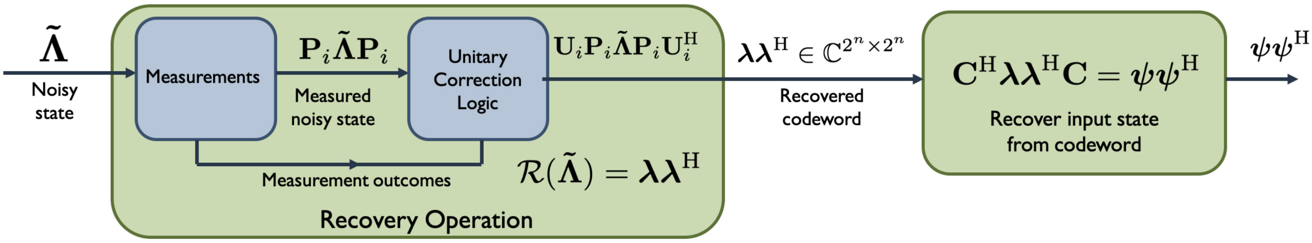
$\tilde{\mathbf{\Lambda}}$
Noisy state
Measurements
$\mathbf{P}_i\tilde{\mathbf{\Lambda}}\mathbf{P}_i$
Measured noisy state
Unitary Correction Logic
$\mathbf{U}_i\mathbf{P}_i\tilde{\mathbf{\Lambda}}\mathbf{P}_i\mathbf{U}_i^{\mathrm{H}}$
Measurement outcomes
$\mathcal{R}(\tilde{\mathbf{\Lambda}}) = \boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{H}}$
Recovery Operation
$\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{H}} \in \mathbb{C}^{2^n \times 2^n}$
Recovered codeword
$\mathbf{C}^{\mathrm{H}}\boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{H}}\mathbf{C} = \boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{H}}$
Recover input state from codeword
$\boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{H}}$

FIGURE 2. The decoder operation up close. The input to the decoder is a random quantum state, represented as a density matrix. In QEC, the decoder makes some measurements, and applies some logic conditioned on the measurement outcomes. This process is known as the *recovery operation*. In the case that the operation succeeds, the output of the recovery operation is a pure state equivalent (up to a global phase) to the transmitted codeword state. Through another measurement, the encoded state can be recovered from the codeword.

measurement outcomes in an effort to recover the transmitted codeword. This process is known as the *recovery operation*. The recovery operation is itself a quantum channel and is denoted $\mathcal{R}$. Assuming that the recovery operation manages to recover the transmitted codeword $\boldsymbol{\lambda}$, the encoded state $\boldsymbol{\psi}$ can be recovered (up to a global phase). An exploded view of the receiver is depicted in Fig. 2.

We say that a code defined by $\mathbf{P}_0$ can correct a channel $\mathcal{E}$ if there exists a physically realizable quantum operation $\mathcal{R}$ (the recovery operation), such that $\mathcal{R}(\mathcal{E}(\mathbf{\Lambda})) = \mathbf{\Lambda}$ for all density matrices $\mathbf{\Lambda}$, such that $\mathbf{P}_0\mathbf{\Lambda}\mathbf{P}_0$ [1]. The *quantum error correction conditions* use the error operators $E = \{\mathbf{E}_i\}$ that parametrize $\mathcal{E}$ to describe whether or a channel can be corrected by a code with projector $\mathbf{P}_0$. The recovery operation $\mathcal{R}$ exists if and only if, for some PSD matrix $\mathbf{A}$ [1, Th. 10.1]

$$\mathbf{P}_0\mathbf{E}_q^{\mathrm{H}}\mathbf{E}_p\mathbf{P}_0 = [\mathbf{A}]_{q,p}\mathbf{P}_0. \tag{14}$$

The “if” direction of this statement is sometimes known as the *sufficiency* of the QEC conditions. The proof of the sufficiency criterion is completely constructive. Given the matrices $\mathbf{E}_i$, the matrix $\mathbf{A}$, and the projector for the codespace, the Kraus operators for the recovery operation $\mathcal{R}$ can be constructed explicitly. It turns out that if (14) holds for some set of operators $E = \{\mathbf{E}_i\}$, it also holds for any set of operators $F = \{\mathbf{F}_i\}$ that can be written as linear combinations of the elements of E [1, Th. 10.2]. In Section III, we discuss the extension of (14) to error operators derived from classical channels where the normalization condition (11b) does not hold. It should be noted that $\mathbf{P}_0$ must have $\operatorname{rank}(\mathbf{P}_0) \geq 2$ to reliably transmit any information. If the code is one dimensional, all codewords are the same (up to a global phase factor) and the measurement statistics at the output will be independent of the input. This is further discussed in Section III.

Stabilizer codes are a class of quantum error correcting codes with a decoding procedure analogous to syndrome decoding in classical linear codes [3]. The stabilizer construction makes use of the group structure of $\mathcal{G}_{\mathcal{P}_n}$ to define both the code subspace $\mathbf{P}_0$ and the decoding procedure. The canonical construction relies on representing the channel's error operators as a linear combinations of the group elements. It turns out that this is always possible. Let $\mathcal{E}$ be a channel with error operators $F = \{\mathbf{F}_i\}$. The elements of F can be expressed as linear combinations of the basis matrices $\mathcal{P}_n$. Let $E = \{\mathbf{E}_i\}$ be the smallest subset of $\mathcal{P}_n$ such that we can write $\mathbf{F}_i = \sum_q c_{i,q}\mathbf{E}_q$ for all i. If code can be shown to satisfy (14) for the set of error operators E, it will automatically satisfy (14) for the set of error operators F. For the remainder of this section, we assume, unless otherwise noted, that the channel $\mathcal{E}$ is parameterized by a set of Pauli error operators E such that $E \subset \mathcal{P}_n$. In fact, the channel $\mathcal{E}$'s actual Kraus operators may be linear combinations of the elements of E.

A stabilizer code with a blocklength of n qubits is defined via an n qubit stabilizer group, G_S. The stabilizer group G_S is a commutative subgroup of $\mathcal{G}_{\mathcal{P}_n}$ that does not contain $-\mathbf{I}_{2^n}$. This latter requirement necessitates that $G_S \subset \mathcal{P}_n$. It can be shown that if the order of the subgroup G_S is $|G_S|$, there is a subset of G_S with $\ell \leq \log_2(|G_S|)$ elements that generates G_S under matrix multiplication [1]. Call this set the *stabilizer group generators*, and denote it by S. Care should be taken not to confuse the stabilizer group generators with the generator matrix for a quantum code, as defined in (12). We refer to the former as “stabilizer group generators,” or “group generators,” and the latter as the “generator matrix for the code,” or a “code's generator matrix.” Since the stabilizer group generators are a subset of G_S, it must be that $S \subset \mathcal{P}_n$ and that the elements of S commute with one another. Assume without loss of generality that S contains ℓ elements and is minimal (i.e., removing one of the group generators reduces the order of the generated group). The stabilizer code defined by S is denoted C_S and is the subspace given by

$$C_S = \left\{\mathbf{v} \in \mathbb{C}^n \quad | \quad \mathbf{v} = \mathbf{S}\mathbf{v} \ \forall \ \mathbf{S} \in G_S\right\}. \tag{15}$$

The stabilizer group generators play a role similar to the rows of the parity check matrix in a classical linear block code. Assume a blocklength of n qubits and that G_S is generated by a minimal set of ℓ stabilizer group generators. In this case, it can be shown that C_S is $2^{(n-\ell)}$ dimensional [3] [1, Prop. 10.5]. Put differently, the code encodes $n - \ell$ message qubits per block. In a classical linear block code, the dimension of the code space is determined by the number of linearly independent rows in the parity check matrix, while in a quantum stabilizer code, the dimension of the code space is determined by the number of independent group generators. The stabilizer group generators, again like the row space of a parity check matrix, are used to diagnose and correct errors induced by the channel on the codeword. This process is described in the following paragraphs.

The stabilizer group generators can be used to determine if some set of Pauli error operators, denoted E, are correctable by the stabilizer code C_S [3], [5]. It can be shown that the projection matrix onto the subspace C_S (cf. (15)) can be written in terms of the stabilizer group generators as

$$\mathbf{P}_0 = \prod_{i=0}^{\ell-1} \frac{(\mathbf{I}_{2^n} + \mathbf{S}_i)}{2}. \tag{16}$$

It turns out that the stabilizer group generators can be used to certify that the QEC conditions (14) hold for the codespace projector $\mathbf{P}_0$ and some set of Pauli error operators $E \subset \mathcal{G}_{\mathcal{P}_n}$. Since both $S \subset \mathcal{G}_{\mathcal{P}_n}$ and $E \subset \mathcal{G}_{\mathcal{P}_n}$, the elements of S and E either commute or anti-commute with one another. As it happens, the code can correct the error operators in E if the elements of E have *unique commutation relationships with respect to the stabilizer group generators* [1], [3]. The pattern of commutations and anti-commutations of a particular error operator $\mathbf{E}_p$ with respect to the group generators S is known as the *syndrome* of $\mathbf{E}_p$. For a code C_S, the syndrome of $\mathbf{E}_p \in E$ is concisely described by a length ℓ binary string $s_{C_S}(\mathbf{E}_p)$ where for $r \in \{0, 1, \ldots, \ell - 1\}$

$$\left[s_{C_S}(\mathbf{E}_p)\right]_r = \begin{cases} 0 & \text{if } \mathbf{E}_p \text{ commutes with } \mathbf{S}_r \\ 1 & \text{if } \mathbf{E}_p \text{ anti-commutes with } \mathbf{S}_r. \end{cases} \tag{17}$$

In other words, (14) is guaranteed to hold for $\mathbf{P}_0$ and E if, for any two distinct error operators $\mathbf{E}_i, \mathbf{E}_p \in E$, we have $s_{C_S}(\mathbf{E}_i) \neq s_{C_S}(\mathbf{E}_p)$.

These results demonstrate an approach to design a stabilizer code for a given channel. Namely, one needs to find a set of commuting Pauli operators that do not generate $-\mathbf{I}_{2^n}$ but have unique commutation relationships with respect to the channel's error operators, E. Consider an analogy to classical binary linear block codes–if the channel and code are such that the mapping from error patterns to syndromes is bijective, the input codeword can be recovered exactly given the syndrome. In the classical case, the syndrome is measured at the receiver via a matrix multiplication. If the error pattern can be uniquely inferred via the syndrome, error correction can be accomplished by simply subtracting the error pattern from the received codeword. Likewise, in the quantum setting, the syndrome is measured at the receiver and a correction is applied conditioned on the measured syndrome. A key distinction of the quantum case is that the measurement process itself changes the received quantum state, blurring the lines between "measurement" and "error correction." It is worth mentioning that the requirement for each error to have a unique syndrome is not necessary for (14) to hold. Codes with a unique syndrome for each error operator are known as non-degenerate, as opposed to degenerate codes (cf. [1], [3]). We presently describe the syndrome measurement and error correction process for stabilizer codes under the assumption that each error has a unique syndrome.

Some assumptions help to illuminate the syndrome measurement. With minimal loss of generality, assume that the generating set S contains exactly $\ell = \log_2(|E|) \in \mathbb{N}$ elements.[1] If each error operator in E has a unique syndrome with respect to the stabilizer group generators, we can assume without loss of generality that the syndrome of the error $\mathbf{E}_p$ is equal to the binary representation of p (the choice of indexing the error operators is arbitrary). Let $[p]_i$ be the i^{th} digit in the binary expansion of the integer $p \in [0, 2^\ell - 1]$. The assumption that $[s_{C_S}(\mathbf{E}_p)]_i = [p]_i$ means that

$$\mathbf{S}_i\mathbf{E}_p = \begin{cases} \mathbf{E}_p\mathbf{S}_i, & \text{if } \left[p\right]_i = 0 \\ -\mathbf{E}_p\mathbf{S}_i, & \text{if } \left[p\right]_i = 1. \end{cases} \tag{18}$$

The syndrome measurement itself is a projective measurement in a carefully chosen measurement basis. Define the following set of projection matrices for $z \in [0, 2^\ell - 1]$

$$\mathbf{P}_z = \prod_{i=0}^{\ell-1} \frac{\left(\mathbf{I}_{2^n} + (-1)^{[z]_i}\mathbf{S}_i\right)}{2}. \tag{19}$$

Note that $\mathbf{P}_0$ is a projector onto the codespace, and that this conforms with the definition in (16). Since the stabilizer group generators commute, we have the commutation property

$$\begin{aligned}&\left(\mathbf{I}_{2^n} + (-1)^{[p]_i}\mathbf{S}_i\right)\left(\mathbf{I}_{2^n} + (-1)^{[p]_j}\mathbf{S}_j\right) \\ &\quad = \left(\mathbf{I}_{2^n} + (-1)^{[p]_j}\mathbf{S}_j\right)\left(\mathbf{I}_{2^n} + (-1)^{[p]_i}\mathbf{S}_i\right). \end{aligned} \tag{20}$$

Assume $p \neq q$. For some i it must be that $[p]_i \neq [q]_i$. Fix this i and recall that $S \subset \mathcal{P}_n$ and thus the elements of S are both Hermitian and unitary. This implies that

$$\left(\mathbf{I}_{2^n} + (-1)^{[p]_i}\mathbf{S}_i\right)\left(\mathbf{I}_{2^n} + (-1)^{[q]_i}\mathbf{S}_i\right) = \mathbf{0}. \tag{21}$$

Given (21) and the aforementioned commutativity relationship (20), one can show that $\mathbf{P}_p\mathbf{P}_q = \mathbf{0}$. It turns out that the projectors (19) all have a rank equal to $2^{n-\ell}$ (cf. [1, Prop. 10.5]). Thus since they are mutually orthogonal, they span $\mathbb{C}^{2^n}$, and can be used to define a projective measurement.

Recall that since the encoded state is defined to be in the codespace, we have $\mathbf{\Lambda} = \mathbf{P}_0\mathbf{\Lambda}\mathbf{P}_0$. The noisy encoded state has the density matrix $\mathcal{E}(\mathbf{\Lambda})$. The receiver performs a projective measurement of $\mathcal{E}(\mathbf{\Lambda})$ in the $\{\mathbf{P}_i\}$ basis. The measurement outcome is a random integer $q \in [0, 2^\ell - 1]$. The measured syndrome is the binary expansion of the measurement outcome.[2] The density matrix after measurement is given by $\mathbf{M}_q$ where

$$\mathbf{M}_q \triangleq \frac{\mathbf{P}_q\mathcal{E}(\mathbf{\Lambda})\mathbf{P}_q}{\mathrm{Tr}\left(\mathbf{P}_q\mathcal{E}(\mathbf{\Lambda})\right)}, \tag{22a}$$

[1] This ensures that the number of possible syndromes $2^{|S|}$ exactly equals the number of error operators, $|E|$.

[2] i.e., the syndrome is the binary string $[q]_0[q]_1 \ldots [q]_{\ell-1}$.

which follows from applying (4) to the density matrix $\mathcal{E}(\mathbf{\Lambda})$. Under the present assumption that each error has a syndrome equal to its binary expansion, i.e., $[s_{C_S}(\mathbf{E}_z)]_i = [z]_i$, we now demonstrate that

$$\mathbf{M}_q = \mathbf{E}_q \mathbf{\Lambda} \mathbf{E}_q^{\mathrm{H}}, \tag{22b}$$

holds when the Kraus operators of $\mathcal{E}$ are exactly the Pauli operators E, or even more generally when the Kraus operators are linear combinations of the elements of E. We will make use of the forthcoming identities in subsequent sections.

Under our assumptions, it can be verified via (19) that for any $p \in \{0, 1, \ldots, \ell - 1\}$

$$\mathbf{P}_p \mathbf{E}_p = \mathbf{E}_p \mathbf{P}_0. \tag{23}$$

This follows from direct substitution and (18), i.e., if $[p]_i = 1$ then $\mathbf{E}_p$ anti-commutes with $\mathbf{S}_i$, and if $[p]_i = 0$ the operators commute. Furthermore, it can also be shown that if $p \neq v$ then for some $x \neq 0$ we have

$$\mathbf{P}_p \mathbf{E}_v = \mathbf{E}_v \mathbf{P}_x \tag{24}$$

which follows precisely from the fact that, necessarily, for some $i \in \{0, 1, \ldots, \ell - 1\}$, we have $[s_{C_S}(\mathbf{E}_v)]_i \neq [p]_i$. The identities (23) and (24) can be combined such that for a state with density matrix $\mathbf{\Lambda} = \mathbf{P}_0 \mathbf{\Lambda} \mathbf{P}_0$ in the code we have

$$\mathbf{P}_p \mathbf{E}_r \mathbf{\Lambda} \mathbf{E}_s^{\mathrm{H}} \mathbf{P}_p = \begin{cases} \mathbf{E}_p \mathbf{\Lambda} \mathbf{E}_p^{\mathrm{H}}, & \text{if } p = r = s \\ \mathbf{0}, & \text{otherwise.} \end{cases} \tag{25}$$

The identity (25) can be used to directly verify that (22b) holds for any quantum channel $\mathcal{E}$ whose error operators are linear combinations of E.

Upon observing the syndrome measurement q, the receiver applies the correction operator $\mathbf{E}_q^{\mathrm{H}}$ which rotates $\mathbf{M}_q$ (cf. (22)) back into the code space via

$$\mathbf{E}_q^{\mathrm{H}} \mathbf{M}_q \mathbf{E}_q = \mathbf{\Lambda}, \tag{26}$$

which shows that an encoded state can be recovered exactly. In stabilizer codes, the syndrome measurement identifies which correction operator will rotate the received state's *post measurement* density matrix back to the original codeword. The recovery process of measurement, followed by correction conditional on the measurement outcome, is represented by the quantum operation $\mathcal{R}$, which acts on the noisy state $\mathcal{E}(\mathbf{\Lambda})$ via

$$\mathcal{R}(\mathcal{E}(\mathbf{\Lambda})) = \sum_i \mathbf{E}_i^{\mathrm{H}} \mathbf{P}_i \mathcal{E}(\mathbf{\Lambda}) \mathbf{P}_i \mathbf{E}_i. \tag{27}$$

If $\mathbf{\Lambda}$ is in the code, and the Kraus operators of $\mathcal{E}$ are linear combinations of the correctable Pauli operators E, (25) can be used to directly verify that $\mathcal{R}(\mathcal{E}(\mathbf{\Lambda})) = \mathbf{\Lambda}$.

In the next section, we demonstrate that a generic class of classical affine channels can be modeled as completely positive convex linear maps. While the maps themselves are not necessarily trace preserving (normalized) like quantum channels, they are amenable to an extension of the error QEC conditions and the corresponding recovery operations. We subsequently apply the formalism to the noncoherent MIMO channel and develop space-time codes using the stabilizer formalism.

III. EMULATED QUANTUM CHANNELS AND RECOVERY OPERATIONS

In this section, we consider one of signal processing's most ubiquitous problems, namely, noncoherent communication over linear (more precisely, affine) channels. We consider classical channel models of the form

$$\mathbf{v} = \overline{\mathbf{H}}\mathbf{t} + \mathbf{n}. \tag{28}$$

We assume that the channel matrix $\overline{\mathbf{H}} \in \mathbb{C}^{m \times m}$ and is generally random. We assume that the additive noise is random and zero-mean. We assume that $\mathbf{t}$ is the transmitted codeword randomly chosen from some codebook, denoted $\mathcal{T}$. We assume that $\mathbf{t}$, $\overline{\mathbf{H}}$, and $\mathbf{n}$ are mutually independent. Unless stated otherwise, we presently make no additional assumptions on the distributions of $\overline{\mathbf{H}}$ and $\mathbf{n}$. Our main interest is noncoherent communication over (28); we assume that the receiver has distributional knowledge of $\overline{\mathbf{H}}$ and $\mathbf{n}$, but does not know the instantaneous realizations. We seek to design a codebook $\mathcal{T}$ and a detection procedure such that the transmitted codeword may be recovered from $\mathbf{v}$ at the receiver. This model is relatively general; for example, the indices of $\mathbf{t}$ could represent a transmission over time, and $\overline{\mathbf{H}}$ could be a Toeplitz matrix implementing convolution with an unknown channel impulse response. This model also corresponds to a vectorized channel model used in communications.

We propose to formulate both the aforementioned (classical) codebook design and detection problems as a quantum coding problem. We propose to view the input signal, $\mathbf{t}$, as an encoded quantum state, and model the classical channel as a deterministic completely positive map. Notably, this map depends only on the moments of the random variables $\overline{\mathbf{H}}$ and $\mathbf{n}$, as opposed to their realizations. We describe a generalization of the QEC conditions to the present classical setting and show that if a quantum code can be found such that these conditions hold exactly, it is possible to reliably detect the elements of a codebook $\mathcal{T}$ derived from the quantum code. The detection rule directly parallels the recovery operation in QEC. In the sequel, we adapt the model (28) to noncoherent MIMO communication. The general technique of defining a quantum analogy of a classical channel may also be of independent interest.

A. MODELING THE TRANSMISSION AS AN ENCODED QUANTUM STATE

To be remain consistent with the previous discussion of qubits, we assume that $m = 2^n$ for some $n \in \mathbb{N}$. We propose to model $\mathbf{t}$ as an n qubit "pure" encoded quantum state. While this assumption is necessary to formulate the resulting quantum coding problem in terms of Pauli Group stabilizer codes, stabilizer coding is not the most general setting for QEC. With the exception of the discussion of stabilizer codes, the material in Section II, as well as this section, follows mutatis mutandis for systems with arbitrary dimension m.

An n qubit state is represented by an equivalence class of vectors in $\mathbb{C}^{2^n}$ that are equivalent up to a global phase, namely

$$\left\{\mathbf{t} : \mathbf{t} = e^{j\theta}\mathbf{t}_0,\ \theta \in [0, 2\pi)\right\}, \tag{29}$$

for some representative $\mathbf{t}_0$ with $\|\mathbf{t}_0\|_2 = 1$. We assume that $\mathbf{t}$ is a codeword from the quantum code defined by the projection matrix (cf. Section II-D) $\mathbf{P}_0 \in \mathbb{C}^{2^n \times 2^n}$. This implies that $\mathbf{t} = \mathbf{P}_0\mathbf{t}$. We assume that the code defined by $\mathbf{P}_0$ encodes d qubits into n (where $n \geq d \geq 1$). Under this assumption $\operatorname{rank}(\mathbf{P}_0) = 2^d$. Recall from (12) the definition of a generator matrix for the quantum code defined by $\mathbf{P}_0$. Let $\mathbf{C} \in \mathbb{C}^{2^n \times 2^d}$ be any such generator matrix for the code. By definition, $\mathbf{C}$ is tall a tall, semi-unitary matrix whose columns span the code subspace. By the orthogonality of the columns we have $\mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}_{2^d}$, and by definition $\mathbf{P}_0 = \mathbf{C}\mathbf{C}^{\mathrm{H}}$. A low-dimensional "message" state $\mathbf{s} \in \mathbb{C}^{2^d}$ is encoded into the higher-dimensional $\mathbf{t} \in \mathbb{C}^{2^n}$ via

$$\mathbf{t} = \sqrt{P_{\mathrm{TX}}}\mathbf{C}\mathbf{s}. \tag{30}$$

We assume $\|\mathbf{s}\|_2 = 1$, and use the factor of $\sqrt{P_{\mathrm{TX}}}$ to normalize the transmit power.

The first step in our proposed modulation is to map bits to some constellation $\mathcal{C}$ of vector symbols $\mathbf{s} \in \mathbb{C}^{2^d}$ with $\|\mathbf{s}\|_2 = 1$ which we think of as representing quantum states to be encoded. Substituting this into (52) gives

$$\mathbf{v} = \overline{\mathbf{H}}\mathbf{C}\left(\sqrt{P_{\mathrm{TX}}}\mathbf{s}\right) + \mathbf{n}, \tag{31}$$

where we now consider $\mathbf{s}$ as the information we would like to recover at the receiver, represented as d qubits. We call $\mathbf{s}$ the transmitted *symbol* corresponding to the codeword $\mathbf{t}$.

Since the symbols $\mathbf{s}$ are themselves to represent quantum states, they are defined only up to a global phase. Thus, an important constraint imposed on the set $\mathcal{C}$ is that for two distinct $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{C}$ and any $\theta \in [0, 2\pi)$

$$\mathbf{s}_1 \neq e^{j\theta}\mathbf{s}_2. \tag{32}$$

The set of d qubit pure states is likewise isomorphic to the set of one-dimensional subspaces of the 2^d dimensional state space, or equivalently a representation of a point on the Grassmann manifold $\mathcal{G}(2^d, 1)$. While we refer to the vector $\mathbf{s}$ as the transmitted symbol, we note that the "information" contained in a symbol is the subspace spanned by $\mathbf{s}$. We propose to design a classical codebook $\mathcal{T}$ using vectors that correspond to the quantum codewords for the symbols. We define

$$\mathcal{T} = \left\{\mathbf{t} | \mathbf{t} = \sqrt{P_{\mathrm{TX}}}\mathbf{C}\mathbf{s},\ \text{where } \mathbf{s} \in \mathcal{C}\right\}. \tag{33}$$

It should be emphasized that the classical code $\mathcal{T}$ is parameterized by both the generator matrix for the quantum code, $\mathbf{C}$, as well as the constellation $\mathcal{C}$. The rate of the classical code necessarily depends on the size of the constellation $\mathcal{C}$. Generally speaking, the design of the constellation $\mathcal{C}$ must also be specified. This is discussed in subsequent sections.

It is worth emphasizing the following particular implication of viewing the transmitted symbol $\mathbf{s}$ as a quantum state. Given a fixed design of the quantum code (in other words, having fixed the code's generator matrix $\mathbf{C}$), the receiver's goal is to recover an estimate of the transmitted symbol $\hat{\mathbf{s}}$ such that $\hat{\mathbf{s}}\hat{\mathbf{s}}^{\mathrm{H}} \approx \mathbf{s}\mathbf{s}^{\mathrm{H}}$.

B. MODELING THE CLASSICAL CHANNEL AS A QUANTUM CHANNEL

Having modeled the transmitted codeword $\mathbf{t}$ as an encoded quantum state, it remains to model the classical channel (28) as a quantum channel (cf. Section II-C). In deference to canonical derivations of quantum channels (cf. [1, Ch. 8.2], [2, Ch. 4]), a reasonable choice for an "emulated" quantum channel $\mathcal{E}$ is to define $\mathcal{E}(\mathbf{t}\mathbf{t}^{\mathrm{H}}) = \mathbb{E}[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}]$. It turns out that this is possible for affine channel models like (28) in the case where the noise is zero-mean and the noise, channel matrix, and transmit codeword are mutually independent, as detailed in the following theorem.

Theorem 1: For a channel model of the form (28) where $\overline{\mathbf{H}} \in \mathbb{C}^{m\times m}$, $\mathbb{E}[\overline{\mathbf{H}}] = \overline{\mathbf{H}}_\mu$ $\mathbb{E}[\mathbf{n}] = \mathbf{0}$, $\mathbb{E}[\mathbf{n}\mathbf{n}^{\mathrm{H}}] = \mathbf{\Phi}$, and $\mathbf{n}$, $\overline{\mathbf{H}}$, and $\mathbf{t}$ are mutually independent, there exists a set of matrices $\mathbf{E}_i \in \mathbb{C}^{m\times m}$ such that

$$\mathbb{E}\left[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}\right] = \sum_i \mathbf{E}_i\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{E}_i^{\mathrm{H}}. \tag{34}$$

We then call the function $\mathcal{E}(\mathbf{t}\mathbf{t}^{\mathrm{H}}) = \mathbb{E}[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}]$ the emulated quantum channel, and we call the matrices $\mathbf{E}_i$ the channel's emulated *Kraus* or *error* operators. The map $\mathcal{E}$ is a convex linear and completely positive function of $\mathbf{t}\mathbf{t}^{\mathrm{H}}$.

A constructive proof of Theorem 1 is provided in Appendix A. The construction leads to the following corollary which says that if the channel is noiseless (i.e., $\mathbf{n} = 0$) and the channel matrix $\overline{\mathbf{H}}$ has all of its energy in a lower-dimensional subspace of $\mathbb{C}^{m\times m}$, then the channel's emulated error operators lie in the same subspace.

Corollary 1: Let $\mathbf{v}$, $\overline{\mathbf{H}}$, and $\mathbf{t}$ be as defined in Theorem 1. Assume $\mathbf{n} = \mathbf{0}$. From Theorem 1, the emulated quantum channel for the classical channel can be written in terms of emulated Kraus operators, e.g., there exists a set of matrices $\{\mathbf{E}_i\}$ such that

$$\mathbb{E}\left[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}\right] \overset{\text{a.s}}{=} \sum_i \mathbf{E}_i\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{E}_i^{\mathrm{H}}. \tag{35}$$

Let $\{\mathbf{\Theta}_0, \ldots, \mathbf{\Theta}_{m^2-1}\}$ denote an arbitrary, complete, orthogonal basis for $m \times m$ matrices. Assume that for some $z \leq m^2 - 1$,

$$\operatorname{Tr}\left(\mathbf{\Theta}_i^{\mathrm{H}}\overline{\mathbf{H}}\right) \overset{\text{a.s}}{=} 0\ \forall\ z \leq i \leq m^2 - 1. \tag{36}$$

Then, without loss of generality all of the error operators in the emulated quantum channel corresponding to the classical channel $\mathbf{v} = \overline{\mathbf{H}}\mathbf{t}$ lie in the span of $\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots \mathbf{\Theta}_{z-1}$. In other words, there exists a set of scalars $\{s_{i,j}\}$ such that for all i

$$\mathbf{E}_i = \sum_{j=0}^{z-1} s_{i,j}\boldsymbol{\Theta}_j. \tag{37}$$

This corollary proves useful in the subsequent discussion of MIMO wireless communication and is proved in Appendix A.

Recall that, to ensure measurement probabilites are properly normalized, a quantum channel acting on a block of n qubits has $\sum_i \mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i = \mathbf{I}_{2^n}$ [2]. In emulated quantum channels, we make no such requirement. Indeed, we generally expect that for the $\mathbf{E}_i$ as defined in Theorem 1, we will have $\sum_i \mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i \neq \mathbf{I}_{2^n}$. Despite this, it turns out that if the condition (14) holds for some projection matrix and the emulated Kraus operators in (34), under mild assumptions the transmitted codeword can be recovered at the receiver. These results are presently summarized and allow us to develop a recipe for classical emulation of the recovery operation from QEC.

C. THE EMULATED RECOVERY OPERATION

The "generalization" of the QEC conditions follows immediately from the standard sufficiency proof of the quantum error correction conditions (cf. [1, Th. 10.1]) and its generalization (cf. [1, Th. 10.2]). For completeness, we formally state the result.

Theorem 2 (The QEC Conditions [1, Ths. 10.1, 10.2]): Define the emulated error operators $E = \{\mathbf{E}_0, \mathbf{E}_1, \ldots \mathbf{E}_{e-1}\}$ as in Theorem 1. Note that $|E| = e$. Assume that

$$\sum_{i=0}^{e-1} \mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i \neq \mathbf{0}. \tag{38}$$

Assume that the (nonzero) projector $\mathbf{P}_0$ satisfies the quantum error correction conditions for a PSD matrix $\mathbf{A} \neq \mathbf{0}$, i.e., assume that

$$\mathbf{P}_0\mathbf{E}_i^{\mathrm{H}}\mathbf{E}_j\mathbf{P}_0 = [\mathbf{A}]_{i,j}\mathbf{P}_0. \tag{39}$$

There exists a set of $r \leq e$ matrices $\mathbf{R}_j$ and a positive constant c such that for all codewords $\mathbf{t} = \mathbf{P}_0\mathbf{t}$ we have

$$\sum_{j=0}^{r-1}\sum_{i=0}^{e-1} \mathbf{R}_j\mathbf{E}_i\mathbf{P}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{P}_0\mathbf{E}_i^{\mathrm{H}}\mathbf{R}_j^{\mathrm{H}} = c\mathbf{P}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{P}_0. \tag{40}$$

Furthermore the matrices $\mathbf{R}_i$ have the form

$$\mathbf{R}_i = \mathbf{U}_i\boldsymbol{\Pi}_i \tag{41}$$

where the $\mathbf{U}_i$ are unitary and the $\boldsymbol{\Pi}_i$ are mutually orthogonal projection matrices such that $\sum_{i=0}^{r-1}\boldsymbol{\Pi}_i = \mathbf{I}_m$.

Moreover, for an arbitrary set of coefficients $\eta_{i,j} \in \mathbb{C}$, with $i, j = \{0, \ldots, e-1\}$, let

$$\mathbf{F}_i = \sum_{j=0}^{e-1} \eta_{i,j}\mathbf{E}_j. \tag{42}$$

Assume that at least one of the $\eta_{i,j}$ is nonzero. If (39) holds for E and $\mathbf{P}_0$, we have, for some positive constant ζ, and the $\mathbf{R}_j$ as defined above (41)

$$\sum_{j=0}^{r-1}\sum_{i=0}^{e-1} \mathbf{R}_j\mathbf{F}_i\mathbf{P}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{P}_0\mathbf{F}_i^{\mathrm{H}}\mathbf{R}_j^{\mathrm{H}} = \zeta\mathbf{P}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{P}_0. \tag{43}$$

While all physically realizable quantum operations (of which quantum channels are a subset) on n qubits are normalized in the sense that $\sum_i \mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i \leq \mathbf{I}_{2^n}$, the sufficiency proof of the quantum error correction conditions (e.g., [1, Th. 10.1]) does not invoke that assumption. Such a condition is only necessary to ensure that measurement probabilities are appropriately normalized and is irrelevant in the present classical setting.

In combination with Theorem 1, Theorem 2 demonstrates that if (for some projector $\mathbf{P}_0$) the error operators in an emulated channel can be shown to satisfy (39), then for a set of matrices $\mathbf{R}_i$ satisfying (41)

$$\sum_{j=0}^{r-1}\sum_{i=0}^{e-1} \mathbf{R}_j\mathbb{E}\big[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}\big]\mathbf{R}_j^{\mathrm{H}} = c\mathbf{P}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{P}_0. \tag{44}$$

Likewise, the conclusion (44) holds if the error operators in an emulated channel are nontrivial linear combinations of some set that satisfy the error correction conditions (39). These results inspire the definition of an "emulated recovery operation" analogous to the recovery operation discussed in Section II-D. Define the matrices $\{\mathbf{R}_i\}$ as in Theorem 2. We define the *emulated quantum recovery operation* as the linear function from PSD matrices to PSD matrices given by

$$f_{\mathcal{R}_C}(\mathbf{W}) = \sum_{j=0}^{r-1} \mathbf{R}_j\mathbf{W}\mathbf{R}_j^{\mathrm{H}}. \tag{45}$$

Consider the random variable $f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}})$. A corollary to the definition of $\mathbf{R}_i$ in Theorem 2 is that if $\mathbb{P}[\mathbf{v} = 0] = 0$, then $f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}}) \overset{\text{a.s}}{\neq} \mathbf{0}$. This observation leads to the final result of this section which completes the generalization of the QEC conditions to the present classical setting.

Theorem 3 (Emulated Recovery Operations): Assume that $\mathbb{E}[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}] = \sum_{i=0}^{e-1}\mathbf{E}_i\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{E}_i^{\mathrm{H}}$, and assume that there exists $\mathbf{P}_0$ such that the conditions of Theorem 2 are satisfied. Let $\mathbf{R}_i$ be as defined in Theorem 2 and define $f_{\mathcal{R}_C}$ as in (45). If $\mathbf{t} = \mathbf{P}_0\mathbf{t}$ and $\mathbf{v} \overset{\text{a.s}}{\neq} \mathbf{0}$, then

$$f_{\mathcal{R}_C}\big(\mathbf{v}\mathbf{v}^{\mathrm{H}}\big) \overset{\text{a.s}}{=} \gamma\mathbf{t}\mathbf{t}^{\mathrm{H}}, \text{ where } \gamma \geq 0 \tag{46}$$

and $\mathbb{P}[\gamma = 0] = 0$.

We defer the proof of this result to Appendix B. Consider a classical channel with input $\mathbf{t}$ and output $\mathbf{v}$. Assume that the channel's emulated error operators $\mathbf{E}_i$ are such that, for some projector $\mathbf{P}_0$, Theorem 2 holds exactly. Further assume that the constraint (32) is satisfied. We now derive a detector to recover the input codeword at the receiver.

Theorem 3 guarantees that computing $f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}})$ at the receiver recovers the subspace spanned by $\mathbf{t}$ (i.e., $\mathbf{t}\mathbf{t}^{\mathrm{H}}$) almost surely. The detection rule

$$\hat{\mathbf{s}} = \underset{\mathbf{s}_i : \mathbf{s}_i \in \mathcal{C}}{\arg\max}\ \mathbf{s}_i^{\mathrm{H}} \mathbf{C}^{\mathrm{H}} f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right) \mathbf{C}\mathbf{s}_i \tag{47}$$

almost surely recovers the input symbol. Recall that for all $\mathbf{s}_i, \mathbf{s}_q \in \mathcal{C}$, $\|\mathbf{s}_i\|_2 = 1$; and that if $i \neq q$, then $\mathbf{s}_i\mathbf{s}_i^{\mathrm{H}} \neq \mathbf{s}_q\mathbf{s}_q^{\mathrm{H}}$. Assume without loss of generality that the transmitted symbol is $\mathbf{s}_i \in \mathcal{C}$. The codeword is then $\mathbf{t} = \mathbf{C}\mathbf{s}_i$. Recall that $\mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}_{2^d}$. By Theorem 3,

$$\mathbf{C}^{\mathrm{H}} f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right)\mathbf{C} \overset{\text{a.s}}{=} \mathbf{s}_i\mathbf{s}_i^{\mathrm{H}}\gamma, \tag{48}$$

and $\gamma \overset{a.s.}{>} 0$. Via Cauchy-Schwartz, for all $\mathbf{u} \in \mathbb{C}^{2^d}$ with $\|\mathbf{u}\|_2 = 1$, we have $|\mathbf{u}^{\mathrm{H}}\mathbf{s}_i|^2 \leq 1$ with equality, if and only if, $\mathbf{u} = e^{j\theta}\mathbf{v}$ for some $\theta \in [0, 2\pi)$. Let $q \neq i$ and $\mathbf{s}_q \in \mathcal{C}$. Since $\gamma \overset{a.s.}{>} 0$ and by construction $\mathbf{s}_q \neq e^{j\theta}\mathbf{s}_i$, we have

$$\mathbf{s}_q^{\mathrm{H}}\mathbf{C}^{\mathrm{H}} f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right)\mathbf{C}\mathbf{s}_q \overset{a.s.}{<} \mathbf{s}_i^{\mathrm{H}}\mathbf{C}^{\mathrm{H}} f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right)\mathbf{C}\mathbf{s}_i. \tag{49}$$

Thus, the detection rule (47) almost surely recovers the input symbol.

Theorems 2 and 3 thus draw an explicit link between quantum error correction and classical noncoherent communication. In Section II-C we noted that, in a quantum communication system, several distinct input-output relationships lead to quantum channels with equivalent Kraus operators. Furthermore, recall from Section II-D that if the quantum error correction conditions hold for a given projector and a set of Kraus operators, they automatically hold for any set of Kraus operators that can be written as linear combinations of the original set. Assume a quantum code exists such that the conditions of Theorem 2 are satisfied for some set of Kraus operators E. Assume that we have a random, classical channel of the form (28), and assume that this channel's emulated quantum channel has Kraus operators that are linear combinations of the elements in E. A system that uses the approach in Section III-A to encode the data at the transmitter and uses the detection rule (47) at the receiver to decode is (almost surely) guaranteed to be error free communication, regardless of the channel's realization.

Recall that the rank of $\mathbf{P}_0$ is the dimension of the quantum state to be encoded. In Section III-A we assumed that $\text{rank}(\mathbf{P_0}) = 2^d$ since d qubits (a 2^d dimensional system) were to be encoded. The symbols $\mathbf{s} \in \mathbb{C}^{2^d}$ were defined to represent pure d qubit pure states. Note that (47) holds irrespective of the cardinality of the constellation $\mathcal{C}$. So, as long as $d \geq 1$, one can choose an arbitrary number of symbols $\mathbf{s}_i$ satisfying (32). Thus, if the modified QEC conditions 2 hold for some emulated quantum channel and have a projector $\mathbf{P}_0$ of at least rank 2, using the detection rule (47) allows finite blocklength reliable communication at arbitrarily high rate.

This demonstrates, in part, that Theorem 3 is a double-edged sword. Consider a classical channel of finite capacity with an emulated quantum channel of the form (34). For the emulated quantum channel, it must be that the conditions of Theorems 2 and 3 cannot hold for a projector $\mathbf{P}_0$ with $\text{rank}(\mathbf{P}_0) \geq 2$.

In the sequel, we develop an emulated quantum channel corresponding to narrowband MIMO communication in Rayleigh fading with AWGN. It can be immediately seen that, for any projector $\mathbf{P}_0$ of at least rank 2, the conditions of Theorems 2 and 3 cannot hold exactly for the derived emulated channel. If such a projector were to exist, the decoding rule based on the emulated recovery operation would allow for reliable, finite blocklength communication at an arbitrary rate. This is, of course, known to be impossible (cf. [38]). It turns out that, in the setting of classical affine channels of the form (28) and their respective emulated quantum channels (cf. Theorem 1), a stronger negative conclusion can be drawn. Theorems 2 and 3 can be used to show, through a similar reductio ad absurdum, that the (modified) QEC conditions in Theorem 2 cannot hold unless the original classical affine channel is noiseless. These aforementioned negative results give insight into quantum error correction and the quantum coding theory in general. We discuss some of these implications in Section VIII.

In many practical settings, quantum error correcting codes have been shown to improve performance in channels where the error correction conditions hold only approximately [1], [5], [39]. Likewise, in classical communication, many noncoherent communication schemes (e.g., [19]) were designed by constructing a code assuming an infinite SNR and then deriving the optimal decoder in the finite SNR case. In the sequel, we proceed along the same lines and design a quantum code for the space-time channel (28) assuming an infinite SNR. We then derive the optimal decoder assuming a Rayleigh fading model with additive white Gaussian noise. In this setting, we demonstrate that the ML decoding rule (at finite SNR) is equivalent to (47). A notable consequence of this approach is that carefully designing the constellation $\mathcal{C}$ can improve performance.

IV. MIMO COMMUNICATION: RECEIVED SIGNAL MODEL AND EMULATED QUANTUM CHANNEL

In this section we present a received signal model for noncoherent narrowband MIMO wireless communication, which is a special case of the model (28). Under some particular assumptions, we use Corollary 1 to show that the error operators of the corresponding emulated quantum channel can be written as a linear combination of Pauli matrices.

A. RECEIVED SIGNAL MODEL

We consider a model analogous to those in [19], [40]. We assume the system has M antennas at both the transmitter and receiver. We assume a narrowband, block fading model with a single-tap (frequency flat) MIMO channel $\mathbf{H} \in \mathbb{C}^{M \times M}$ that is constant over T symbols. In the noncoherent setting, the receiver does not know the realization of $\mathbf{H}$. For the model to be amenable to the design of a stabilizer code, we take $M = 2^k$ and $T = 2^\tau$. Let $\mathbf{T} \in \mathbb{C}^{M \times T}$ be a space-time codeword transmitted over one coherence interval, and let $\mathbf{N}$ be the additive noise. We assume that $\mathbf{N}$ is zero-mean, and that $\mathbf{H}$, $\mathbf{N}$, and $\mathbf{T}$ are mutually independent. We assume $\mathbf{H} \overset{\text{a.s}}{\neq} \mathbf{0}$. The received signal, $\mathbf{V} \in \mathbb{C}^{M\times T}$, is given by

$$\mathbf{V} = \mathbf{HT} + \mathbf{N}. \tag{50}$$

Define the vectorizations $\mathbf{v} = \text{vec}(\mathbf{V}) \in \mathbb{C}^{TM}$, $\mathbf{t} = \text{vec}(\mathbf{T}) \in \mathbb{C}^{TM}$, $\mathbf{n} = \text{vec}(\mathbf{N}) \in \mathbb{C}^{TM}$, and let

$$\overline{\mathbf{H}} = \mathbf{I}_T \otimes \mathbf{H}. \tag{51}$$

The model (50) is equivalent to

$$\mathbf{v} = \overline{\mathbf{H}}\mathbf{t} + \mathbf{n}, \tag{52}$$

which is, as promised, a special case of the model (28).

In the following subsection, we specialize our work in Section III to the space-time channel model. In particular, we will develop an emulated quantum channel model corresponding to (52) at infinite SNR, e.g., when $\mathbf{N} = \mathbf{0}$. We will demonstrate that the error operators in the emulated quantum channel can be written as linear combinations of a subset of the Pauli group $\mathcal{P}_{\tau+k}$. This result does not depend on any particular model for the channel matrix $\mathbf{H}$. To keep the next subsection sufficiently general, for now we only assume that $\mathbf{H}$ and $\mathbf{T}$ are mutually independent. In later sections, we will assume a Rayleigh fading model where the entries of $\mathbf{H}$ as well as the additive noise are complex Gaussian distributed. Rayleigh fading models are appropriate, for example, in a frequency hopping system operating in an environment with rich scattering.

B. THE EMULATED QUANTUM CHANNEL MODEL AT HIGH SNR

Following from the discussion in Section III-A, we propose to use the codewords of a quantum code as vectorized space-time codewords, exactly analogously to the proposed codebook (33). Since $\mathbf{t} \in \mathbb{C}^{MT}$ where $MT = 2^{k+\tau}$, we think of $\mathbf{t}$ as an $n = k + \tau$ qubit pure state. We assume that $\mathbf{t}$ is a codeword of a blocklength n quantum code that encodes d logical qubits. Let $\mathbf{C} \in \mathbb{C}^{2^n\times 2^d}$ denote the quantum code's (tall, semi-unitary) generator matrix and let $\mathcal{C}$ denote a constellation of d qubit pure state vector symbols. The space-time code C_{ST} is given by the set of vectorized codewords

$$C_{ST} = \left\{\mathbf{t}|\mathbf{t} = \sqrt{T}\mathbf{C}\mathbf{s}, \text{ where } \mathbf{s} \in \mathcal{C}\right\}. \tag{53}$$

The space-time code itself is thus parameterized by both $\mathbf{C}$ and $\mathcal{C}$. Since $\mathbf{C}$ has orthonormal columns and the symbols $\mathbf{s} \in \mathcal{C}$ have $\|\mathbf{s}\|_2 = 1$, for all codewords $\mathbf{t} \in C_{ST}$, $\|\mathbf{t}\|_2 = T$. Thus, the time-averaged radiated power satisfies

$$\frac{\mathbb{E}\left[\|\mathbf{t}\|_2^2\right]}{T} = 1. \tag{54}$$

The rate of the space-time code is determined by the size of the constellation $\mathcal{C}$ and is given by

$$R = \frac{\log_2(|\mathcal{C}|)}{T} \tag{55}$$

in bits/channel use. The quantum coding theory is concerned with the design of $\mathbf{C}$ rather than $\mathcal{C}$, and we design $\mathbf{C}$ assuming that the channel (52) is noiseless. In the noiseless case, the design of the constellation $\mathcal{C}$ is arbitrary (up to the constraint (32)). We choose the constellation $\mathcal{C}$ as a Grassmannian packing in 2^d dimensions, which is later shown to minimize an upper bound on the probability of error in the case of Rayleigh fading with Gaussian noise. Note that while many noncoherent space-time codes are designed via Grassmannian packings of the row space of $\mathbf{T} \in \mathbb{C}^{2^k\times 2^\tau}$ [19], the packing used to design $\mathcal{C}$ occurs in a lower-dimensional ($2^d < 2^{k+\tau}$dimensional) space. In other words, our approach maps these lower-dimensional packings to codewords in a higher-dimensional space, resulting in a constellation of space-time codewords $\mathbf{T}$ parameterized by the lower-dimensional vector $\mathbf{s}$. Note also that this choice of $\mathcal{C}$ is guaranteed to satisfy the constraint (32).

At infinite SNR, the channel model (52) may be written

$$\mathbf{v} = \overline{\mathbf{H}}\mathbf{t}, \tag{56}$$

where $\overline{\mathbf{H}} = \mathbf{I}_T \otimes \mathbf{H}$. The emulated quantum channel corresponding to (56) is defined as

$$\mathcal{E}\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right) = \mathbb{E}\left[\overline{\mathbf{H}}\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}^{\mathrm{H}}|\mathbf{t}\right]. \tag{57}$$

By Theorem 1, $\mathcal{E}$ has a Kraus operator representation of the form (34). Rather than compute the operators explicitly, we use Corollary 1 to demonstrate that the error operators can be written as linear combinations of a subset of the Pauli group. By definition, $T = 2^\tau$ and $M = 2^k$, so $\overline{\mathbf{H}} \in \mathbb{C}^{2^{k+\tau}\times 2^{k+\tau}}$ has $\overline{\mathbf{H}} = \mathbf{I}_{2^\tau} \otimes \mathbf{H}$. The subset $\mathcal{P}_{k+\tau}$ of the $k+\tau$ qubit Pauli group is a complete and orthogonal basis for $\mathbb{C}^{2^{k+\tau}\times 2^{k+\tau}}$ (cf. Section II-B). Any arbitrary $\mathbf{J} \in \mathcal{P}_{k+\tau}$ can be written as the Kronecker product of some $\mathbf{J}_\tau \in \mathcal{P}_\tau$ and $\mathbf{J}_k \in \mathcal{P}_k$ via

$$\mathbf{J} = \mathbf{J}_\tau \otimes \mathbf{J}_k. \tag{58}$$

Consider the inner product between $\mathbf{J}$ and $\overline{\mathbf{H}}$

$$\text{Tr}\left(\mathbf{J}^{\mathrm{H}}\overline{\mathbf{H}}\right) = \text{Tr}\left(\left(\mathbf{J}_\tau^{\mathrm{H}} \otimes \mathbf{J}_k^{\mathrm{H}}\right)\left(\mathbf{I}_{2^\tau} \otimes \mathbf{H}\right)\right), \tag{59a}$$

where we have used (51) and (58). From the bilinearity of the Kronecker product, we have that $(\mathbf{J}_\tau^{\mathrm{H}} \otimes \mathbf{J}_k^{\mathrm{H}})(\mathbf{I}_{2^\tau} \otimes \mathbf{H}) = (\mathbf{J}_\tau^{\mathrm{H}}) \otimes (\mathbf{J}_k^{\mathrm{H}}\mathbf{H})$. If $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$, we have $\text{Tr}(\mathbf{C}) = \text{Tr}(\mathbf{A})\text{Tr}(\mathbf{B})$. Applying this identity yields

$$\text{Tr}\left(\mathbf{J}^{\mathrm{H}}\overline{\mathbf{H}}\right) = \text{Tr}\left(\mathbf{J}_\tau^{\mathrm{H}}\right)\text{Tr}\left(\mathbf{J}_k^{\mathrm{H}}\mathbf{H}\right). \tag{59b}$$

Recall that $\mathbf{I}_{2^\tau} \in \mathcal{P}_\tau$. The orthogonality of the matrices in $\mathcal{P}_\tau$ guarantees that $\text{Tr}(\mathbf{J}_\tau^{\mathrm{H}}) = 0$, unless $\mathbf{J}_\tau = \mathbf{I}_{2^\tau}$. Consider an arbitrary enumeration of the 4^k elements in $\mathcal{P}_k$ so that $\mathcal{P}_k = \{\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_{4^k-1}\}$. Define the set of matrices $E = \{\mathbf{E}_i\}$ via

$$\mathbf{E}_i = \mathbf{I}_{2^\tau} \otimes \mathbf{R}_i \quad \text{where } i \in \left\{0, 1, \ldots, 4^k - 1\right\}. \tag{60}$$

It follows that $\overline{\mathbf{H}}$ can be written as a linear combination of the basis matrices (60). This is unsurprising, given that there are precisely 4^k orthogonal matrices in E, and that the arbitrary channel matrix has $\mathbf{H} \in \mathbb{C}^{2^k \times 2^k}$. Let $F = \{\mathbf{F}_i\}$ denote the set of error operators that parametrize the emulated quantum channel corresponding to (56). By (59b) and Corollary 1, we have that each $\mathbf{F}_i \in F$ can be written as a random linear combination of the matrices in $\mathbf{E}_i \in E$. In other words, irrespective of the distribution of $\mathbf{H}$, for a set of coefficients $\{d_{i,j}\}$ we have

$$\mathbf{F}_i = \sum_{j=0}^{4^k-1} d_{i,j} \mathbf{I}_{2^\tau} \otimes \mathbf{R}_j. \tag{61}$$

Recall that if a quantum code can correct some set of errors, it can also correct linear combinations of them (cf. Theorem 3). We conclude that any quantum code (encoding at least a single qubit) that can correct the Pauli errors in E can be used to reliably communicate over the (noiseless) classical channel (56). This observation holds irrespective of the distribution of $\mathbf{H}$, so long as $\mathbf{H} \overset{\text{a.s}}{\neq} \mathbf{0}$. We propose to choose $\mathbf{C}$ as the generator matrix of a quantum code that can correct the error operators in E.

While it was specified that $M = 2^k$, $T = 2^\tau$, and that the quantum code with generator $\mathbf{C}$ encodes a d qubit message into a block of $n = k + \tau$ qubits, we have yet to specify any specific relationship between k, τ, and d. Fixing the number of antennas, the blocklength of the quantum code, n, is controlled by the classical channel's coherence time $T = 2^\tau$. The number of elements in E is precisely the number of elements in $\mathcal{P}_k$ (cf. (60)), namely $|E| = 4^k = M^2$. Recall that a non-degenerate stabilizer code has a unique syndrome for each error operator in E. It can be shown that a code that encodes d logical qubits of information and perfectly corrects $|E|$ errors without degeneracy must have a blocklength n satisfying $n \geq \log_2(2^d |E|)$ [1]. We insist that at least one qubit be encoded (e.g., $d \geq 1$), which ensures that the quantum code is a least 2 dimensional. Applying the definitions directly gives

$$k + \tau \geq d + 2k. \tag{62}$$

This implies that $2^\tau \geq 2^{d+k}$. Since $d \geq 1$, this implies that the coherence time T be such that

$$T \geq 2M. \tag{63}$$

While this bound followed purely from analyzing a quantum coding problem, it can be shown that, for a general noncoherent space-time block code, having $T \geq 2M$ is required to obtain full diversity (and thus competitive performance at a high SNR) [36]. Furthermore, (63) is also reminiscent of information theoretic arguments suggesting that at high SNR, at most $\lfloor T/2 \rfloor$ transmit antennas can be used to achieve capacity [19]. Given these observations, as well as our focus on low-latency communication, we set $d = 1$ and $\tau = k + 1$ so that $T = 2M$. This assumption also expedites comparisons with popular coherent schemes; M time instances can be used for channel estimation, followed by M uses for communication. Notably, the requirement that the quantum code be non-degenerate is not necessary. With degeneracy it is possible to extended the space-time code we propose to the case of $T = M = 2$, although the results were underwhelming. This is discussed in Section VIII.

In the next section, assuming $M = 2^k$ and $T = 2^{k+1}$, we design a quantum code that encodes 1 qubit into an $n = 2k + 1$ qubit block. The code can correct the 4^k errors in the set E (as defined in (60)). As detailed in Section III, the stabilizer code and its emulated recovery operation, can be used to achieve reliable communication over the noiseless channel (56). In the following section, we analyze the space-time code in the special case of Rayleigh fading at finite SNR, where it is shown to exhibit competitive performance.

V. A STABILIZER CODE FOR THE SPACE-TIME CHANNEL

In this section we design a stabilizer code for the noiseless space-time channel. We motivate our construction by explicitly constructing a space-time code for $M = 2$ antennas in the next subsection. While this derivation follows that of our previous paper [37], the space-time codes derived here can be more expediently generalized to larger antenna arrays, namely those where $M = 2^k$. We discuss this recursive generalization in Section V-B.

A. THE 2 × 2 CASE

In keeping with the notation of Section IV, we assume $k = 1$ which gives $M = 2$ and $T = 4$. As proposed in Section IV-B, we seek to design a quantum code that encodes $d = 1$ qubit into $n = 3$ qubits and that can correct the error operators E defined in (60). Explicitly, $E = \{\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ where

$$\mathbf{E}_0 = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{I}_2 \tag{64a}$$

$$\mathbf{E}_1 = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{X} \tag{64b}$$

$$\mathbf{E}_2 = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{Z} \tag{64c}$$

$$\mathbf{E}_3 = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{Y}, \tag{64d}$$

where we have used the fact that $\mathbf{I}_4 = \mathbf{I}_2 \otimes \mathbf{I}_2$. Since $E \subset \mathcal{P}_n$, all of the $\mathbf{E}_i$ are both Hermitian and unitary.

Given that the error operators in E are Pauli operators, stabilizer codes are an attractive approach to the quantum code design. Recall from Section II-D that a stabilizer code that encodes d qubits into n is parameterized by a set of $\ell = n - d$ stabilizer group generators. In the present case, we have $\ell = 2$. Denote the set of stabilizer group generators $S = \{\mathbf{S}_0, \mathbf{S}_1\}$. In particular, given the assumption that $T = 2M$ (cf. (62) and (63)), we can require that the stabilizer code be nondegenerate. Recall that the number of possible syndromes is equal to 2^ℓ. Under the present assumptions, the number of possible syndromes (2^2) is exactly equal to the number of error operators in E.

Recall that $S \subset \mathcal{P}_3$, and that the elements of S must commute and not generate $-\mathbf{I}_{2^3}$. Finally, the stabilizer code generated by S is guaranteed to satisfy the QEC conditions (39) if the elements of S have unique commutation relationships, or syndromes, with respect to the elements

TABLE 1. **Summary of commutation relations between stabilizer and error operators. 0 denotes commutation and 1 denotes anti-commutation.**

Commutation Relationships		
	$\mathbf{S}_0$	$\mathbf{S}_1$
$\mathbf{E}_0 = \mathbf{I}_4 \otimes \mathbf{I}_2$	0	0
$\mathbf{E}_1 = \mathbf{I}_4 \otimes \mathbf{X}$	0	1
$\mathbf{E}_2 = \mathbf{I}_4 \otimes \mathbf{Z}$	1	0
$\mathbf{E}_3 = \mathbf{I}_4 \otimes \mathbf{Y}$	1	1

of E. Two commuting Pauli operators with unique commutation relationships with respect to the elements of E are

$$\mathbf{S}_0 = \mathbf{X} \otimes \mathbf{X} \otimes \mathbf{X} \text{ and} \tag{65a}$$

$$\mathbf{S}_1 = \mathbf{X} \otimes \mathbf{Z} \otimes \mathbf{Z}. \tag{65b}$$

The fact that the elements of S (and thus the subgroup they generate) are commutative can be verified by applying (9) to $\mathbf{S}_0$ and $\mathbf{S}_1$. Furthermore, it can be shown explicitly that $\mathbf{S}_0^i \mathbf{S}_1^p \neq -\mathbf{I}_8$ for all i, p. It is clear that $\mathbf{S}_0 \neq \mathbf{S}_1^i$ for any i, and that $\mathbf{S}_0^i \neq \mathbf{S}_1$, so the set of group generators is minimal. The set of commutation relations, or syndrome, of error $\mathbf{E}_i \in E$ is the binary expansion along the rows of the i^{th} row of Table 1. Note that by construction, the syndrome for error $\mathbf{E}_i$ is the (big-endian) binary expansion of i.

To devise the encoding and recovery operations, first define the following set of orthogonal projectors

$$\mathbf{P}_0 = \frac{(\mathbf{I}_2 + \mathbf{S}_0)(\mathbf{I}_2 + \mathbf{S}_1)}{4} \tag{66a}$$

$$\mathbf{P}_1 = \frac{(\mathbf{I}_2 + \mathbf{S}_0)(\mathbf{I}_2 - \mathbf{S}_1)}{4} \tag{66b}$$

$$\mathbf{P}_2 = \frac{(\mathbf{I}_2 - \mathbf{S}_0)(\mathbf{I}_2 + \mathbf{S}_1)}{4} \tag{66c}$$

$$\mathbf{P}_3 = \frac{(\mathbf{I}_2 - \mathbf{S}_0)(\mathbf{I}_2 - \mathbf{S}_1)}{4}. \tag{66d}$$

By construction (cf. (19) and [1, Prop. 10.5]) this set of projectors are mutually orthogonal and each spans a 2 dimensional subspace. Recall that by definition, $\mathbf{P}_0$ is a projector onto the quantum code's subspace. The generator matrix for the code $\mathbf{C} \in \mathbb{C}^{8\times 2}$ is an arbitrary semi-unitary matrix satisfying

$$\mathbf{C}\mathbf{C}^{\mathrm{H}} = \mathbf{P}_0. \tag{67}$$

As shown in Section II-D, the stabilizer formalism provides the following convenient recipe for the (emulated) recovery operation

$$f_{\mathcal{R}_C}\big(\mathbf{v}\mathbf{v}^{\mathrm{H}}\big) = \sum_{i=0}^{3} \mathbf{E}_i^{\mathrm{H}} \mathbf{P}_i \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{P}_i \mathbf{E}_i. \tag{68}$$

So long as $\mathbf{H} \overset{\text{a.s}}{\neq} \mathbf{0}$, if $\mathbf{v}$ is defined via (56) and $\mathbf{t} = \sqrt{T}\mathbf{C}\mathbf{s}$ for some $\mathbf{s} \in \mathbb{C}^2$, the Theorem 3 guarantees that for some real random variable $c \overset{\text{a.s}}{>} 0$

$$f_{\mathcal{R}_C}\big(\mathbf{v}\mathbf{v}^{\mathrm{H}}\big) = c\mathbf{t}\mathbf{t}^{\mathrm{H}}. \tag{69}$$

In the sequel, we will demonstrate that (69) arises naturally in the derivation of the maximum likelihood decoder for the Rayleigh fading case. In the remainder of this section, we discuss a generalization of the stabilizer code defined by (65) to $M = 2^k$ and $T = 2^{k+1}$ for a general $k = 2, 3, \ldots$

B. THE MORE GENERAL CASE OF M = 2^K ANTENNAS

Having explicitly constructed a code for the case $M = 2$ and $T = 4$, we generalize the construction for $M = 2^k$ and $T = 2^{k+1}$. For a general k, we seek to design a quantum stabilizer code that encodes 1 qubit into $n = 2k + 1$ qubits (cf. Section IV). The channel $\mathbf{H}$ is a $2^k \times 2^k$ matrix. Thus, to guarantee perfect performance at an infinite SNR for arbitrary $\mathbf{H}$, the quantum code must correct the $M^2 = 4^k$ error operators in the set E^k defined by (cf. (60))

$$E^k = \big\{\mathbf{I}_{2^{k+1}} \otimes \mathbf{R} : \mathbf{R} \in \mathcal{P}_k\big\}. \tag{70}$$

As demonstrated in Section IV-B, E^k contains $M^2 = 4^k$ elements. In this section, we construct a code that is (in some sense) a recursive generalization of the code constructed in Section V-A.

One of the most challenging aspects of verifying that the generalization actually works is keeping track of syndromes. To simplify subsequent calculations, we enumerate the elements of $\mathcal{P}_k$ according to a specific choice of indexing. Formally, this choice of indexing is a bijection $\mathbf{\Gamma}_k : \{0, 1, \ldots, 4^k - 1\} \to \mathcal{P}_k$. Consider first the bijection $\mathbf{\Omega} : \{0, 1, 2, 3\} \to \mathcal{P}_1$:

$$\mathbf{\Omega}(q) = \begin{cases} \mathbf{I}_2 & q = 0 \\ \mathbf{X} & q = 1 \\ \mathbf{Z} & q = 2 \\ \mathbf{Y} & q = 3. \end{cases} \tag{71}$$

Each $z \in \{0, 1, \ldots, 4^k - 1\}$ has a unique k digit 4-ary expansion, which we denote $(z_0, z_1, \ldots, z_{k-2}, z_{k-1})_4$. We consider z_{k-1} to be the least significant digit (i.e., big-endian notation where the most significant term, z_0, has the lowest address). Consider the bijection $\mathbf{\Gamma}_k$:$\{0, 1, \ldots, 4^k - 1\} \to \mathcal{P}_k$:

$$\mathbf{\Gamma}_k(z) = \mathbf{\Omega}(z_0) \otimes \mathbf{\Omega}(z_1) \otimes \ldots \otimes \mathbf{\Omega}(z_{k-2}) \otimes \mathbf{\Omega}(z_{k-1}). \tag{72}$$

We enumerate the error operators $E^k = \{\mathbf{E}_0^k, \ldots, \mathbf{E}_{4^k-1}^k\}$ by

$$\mathbf{E}_i^k = \mathbf{I}_{2^{k+1}} \otimes \mathbf{\Gamma}_k(i), \tag{73}$$

where $i \in \{0, 1, \ldots, 4^k - 1\}$.

For the quantum code to encode a single logical qubit (equivalently, for the subspace defining the code to be 2-dimensional), there must be exactly $\ell = n - 1 = 2k$ independent stabilizer group generators (see the discussion after (15)). Note that the number of possible syndromes is exactly $2^\ell = 4^k = M^2$, which is equal to the number of error operators in E^k. Thus, we seek a non-degenerate code where each error operator in E^k has a unique syndrome with respect to the stabilizer group generators. While, in principle, there are several possible sets of stabilizer group generators that generate a code of the form (15) that are guaranteed to correct the errors in E^k, we propose a particular set of group generators that is a recursive generalization of (65). For $i \in \{0, 1, \dots, 2k-1\}$, define the stabilizer group generators as

$$\mathbf{S}_i^k = \begin{cases} \mathbf{X} \otimes \mathbf{I}_{2^{\lfloor \frac{i}{2} \rfloor}} \otimes \mathbf{X} \otimes \mathbf{I}_{2^{k-1}} \otimes \mathbf{X} \otimes \mathbf{I}_{2^{k-1-\lfloor \frac{i}{2} \rfloor}}, & i \text{ even}, \\ \mathbf{X} \otimes \mathbf{I}_{2^{\lfloor \frac{i}{2} \rfloor}} \otimes \mathbf{Z} \otimes \mathbf{I}_{2^{k-1}} \otimes \mathbf{Z} \otimes \mathbf{I}_{2^{k-1-\lfloor \frac{i}{2} \rfloor}}, & i \text{ odd}. \end{cases} \tag{74}$$

Note that choosing $k = 1$ gives the same stabilizer group generators as (65).

To show that the stabilizer group generators $S^k = \{\mathbf{S}_0^k, \mathbf{S}_1^k, \dots\}$ defined by (74) generate a stabilizer code, we must show that (for all $k \in \mathbb{N}$), the operators in S^k are mutually commuting, do not generate $-\mathbf{I}_{2^{2k+1}}$, and are in fact minimal. We outline how to verify this in the theorem below.

Theorem 4 (For All k, S^k is a Minimal Set of Stabilizer Group Generators): Let $S^k/\mathbf{S}_i^k$ denote the set S^k with the element $\mathbf{S}_i^k$ removed. For all k, we have that:

1) The matrices in S^k commute with one another.
2) No product of matrices in S^k is equal to $-\mathbf{I}_{2^{2k+1}}$.
3) No product of the matrices in $S^k/\mathbf{S}_i^k$ can generate $\mathbf{S}_i^k$.

The first and second of these conclusions guarantee that the subgroup generated by S^k is commutative and does not contain $-\mathbf{I}_{2^{2k+1}}$, which in turn demonstrates that the subgroup generated by S^k is a valid stabilizer group. The third conclusion guarantees that the set of stabilizer group generators is minimal.

Proof: Note that $S^k \subset \mathcal{P}_{2k+1}$. Commutation is easiest to verify when armed with the properties of the Pauli group. Begin by writing the matrices in (74) as a product of $2k+1$ operators in $\mathcal{P}_1$; all that is required is to expand the identity matrices into products of $\mathbf{I}_2$ (e.g., $\mathbf{I}_{2^{k-1}} = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \cdots \otimes \mathbf{I}_2$). Appropriating the terminology from quantum information, we say that the matrices $\mathbf{S}_i^k$ operate on $2k+1$ qubits. Recall from Section II-B that two operators in $\mathcal{P}_{2k+1}$ commute if and only if the number of qubits on which the operators perform different non-identity transformations is even. Otherwise, they anti-commute.

Consider $\mathbf{S}_q^k$ and $\mathbf{S}_r^k$. All operators in (74) perform the same $\mathbf{X}$ transformation on the first (leftmost) qubit. If $\lfloor \frac{q}{2} \rfloor \neq \lfloor \frac{r}{2} \rfloor$, then for all but the first qubit (where they perform the same operation), at least one of the operators performs an $\mathbf{I}_2$ transformation. Thus the operators commute. If $\lfloor \frac{q}{2} \rfloor = \lfloor \frac{r}{2} \rfloor$, it suffices to consider the case that q is even and $r = q + 1$. It immediately follows from (74) that non-identity transformations are computed only on the first, $(2 + q/2)^{\text{th}}$, and $(3 + k - 1 + q/2)^{\text{th}}$ qubits. In both operators, an $\mathbf{X}$ is performed on the first qubit. By construction, $\mathbf{S}_q$ performs an $\mathbf{X}$ on the $(2 + q/2)^{\text{th}}$ and $(3 + k - 1 + q/2)^{\text{th}}$ qubits while $\mathbf{S}_r$ performs a $\mathbf{Z}$ on those qubits. Since the number of differences is even, the operators commute.

A proof that the elements of S^k fail to generate $-\mathbf{I}_{2^{2k+1}}$ follows from an induction on k. The base case for $k = 1$ follows from checking by "brute force" that the matrices in (65) cannot generate $-\mathbf{I}_8$. The induction follows by recognizing the recursive relationship between S^n and S^{n+1}. In particular, for $i > 1$, for some $\mathbf{W}_i \in \mathcal{P}_n$, the stabilizer group generator $\mathbf{S}_{i-2}^n$ has the form (cf. (74))

$$\mathbf{S}_{i-2}^n = \mathbf{X} \otimes \mathbf{W}_i \otimes \mathbf{W}_i. \tag{75a}$$

For the same operator $\mathbf{W}_i$, the stabilizer group generator $\mathbf{S}_i^{n+1} \in S^{n+1}$ has the form (cf. (74))

$$\mathbf{S}_i^{n+1} = \mathbf{X} \otimes \mathbf{I}_2 \otimes \mathbf{W}_i \otimes \mathbf{I}_2 \otimes \mathbf{W}_i. \tag{75b}$$

See Appendix C for the complete proof.

The group generators defined by (74) can be shown to be minimal by noting that if one removes $\mathbf{S}_i^k$ from S^k, one cannot generate $\mathbf{S}_i^k$. One can show that for all arbitrary $z_0, z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_{2k-1} \in \mathbb{Z}$, $\mathbf{S}_i^k \neq (\mathbf{S}_0^k)^{z_0}(\mathbf{S}_1^k)^{z_1} \dots (\mathbf{S}_{z_{i-1}}^k)^{z_{i-1}}(\mathbf{S}_{z_{i+1}}^k)^{z_{i+1}} \dots (\mathbf{S}_{2k-1}^k)^{z_{2k-1}}$. Assuming that i is even, it can be shown that without loss of generality $z_j = 0$ for all $j \neq i+1$. It then suffices to show that $\mathbf{S}_{i+1}^{z_{i+1}} \neq \mathbf{S}_i$ for any z_{i+1}, which is follows directly from the base case in Section V-A. An analogous argument can be made for i odd. ■

Theorem 4 guarantees that for $z \in \{0, 1, \dots, 4^k - 1\}$ and $(z_0 z_1 z_2 \dots z_{2k-1})_2$ the big-endian binary expansion of z, the following set of projection matrices are mutually orthogonal and each spans a two-dimensional subspace (cf. Section II-D)

$$\mathbf{P}_z = \prod_{i=0}^{2k-1} \frac{\left(\mathbf{I}_{2^{2k+1}} + (-1)^{z_i}\mathbf{S}_i^k\right)}{2}. \tag{76}$$

Since there are exactly 4^k such 2 dimensional projectors, the orthogonality ensures that they form a complete set (they span $\mathbb{C}^{2^{2k+1} \times 2^{2k+1}}$). The quantum code itself, denoted C_{S^k}, is the subspace defined by $\mathbf{P}_0$. Analogously to (67), we define the code's generator matrix (cf. (30), (53)), $\mathbf{C} \in \mathbb{C}^{2^{2k+1} \times 2}$ as a semi-unitary matrix given by

$$\mathbf{C}\mathbf{C}^{\mathrm{H}} = \mathbf{P}_0. \tag{77}$$

It remains to be shown that C_{S^k} (equivalently $\mathbf{P}_0$) satisfies the QEC conditions (14) for the set of error operators E^k. Again, this can be established by verifying that the error operators in E^k have a unique syndrome with respect to the stabilizer group generators S^k. It turns out, in analogy to Table 1, that the syndrome of error $\mathbf{E}_i \in E^k$ with respect to the operators in S^k is exactly equal to the binary expansion of i, thus guaranteeing that each error has a unique syndrome. The result is summarized formally in the following theorem.

Theorem 5: Define the elements in E^k as in (73) and the stabilizer group generators in S^k as in (74). The syndrome of the error $\mathbf{E}_r^k$, with respect to the stabilizer group generators S^k, is the length $2k$ binary string denoted $s_{C_{S^k}}(\mathbf{E}_i^k)$ and is defined as in (17). For all $k \in \mathbb{N}$ and $r \in \{0, 1, \dots, 4^k - 1\}$, we have that

$$s_{C_{S^k}}\left(\mathbf{E}_r^k\right) \text{ is the big-endian binary expansion of } r. \tag{78}$$

Proof: The proof again follows from induction on k with the $k = 1$ base case as given in Table 1. The proof follows from the recurrence relationship between stabilizer group generators S^n and S^{n+1} (i.e., (75)) as well as a recurrence relationship between the error operators E^n and E^{n+1}. The remaining details of the proof are given in Appendix D. ■

The recovery operation follows from the formula in (27), via

$$f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right) = \sum_{i=0}^{4^k-1} \mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right)\mathbf{P}_i\mathbf{E}_i. \tag{79}$$

Let $C_{ST}(k)$ be a space-time code as defined in (53) with the matrix $\mathbf{C}$ given by (77). Let $\mathbf{v}$ be the received signal at infinite SNR (cf. (56)). So long as $\overline{\mathbf{H}} \overset{\text{a.s}}{\neq} 0$, if $\mathbf{t}$ is a codeword of C_{ST} we have

$$f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right) = c\mathbf{t}\mathbf{t}^{\mathrm{H}}, \tag{80}$$

where $c \overset{\text{a.s}}{>} 0$. In the next section, we demonstrate that in the special case of Rayleigh fading and additive white Gaussian noise, (79) is a sufficient statistic for maximum likelihood (ML) detection of the symbols encoded via (30). In fact, the emulated recovery operation is a natural design for an ML receiver.

VI. MAXIMUM LIKELIHOOD (ML) DECODING FOR NOISY CHANNELS

In the remainder of this manuscript, we focus on the application of the previous section's space-time code construction to the canonical setting of Rayleigh fading with additive white Gaussian noise. In Rayleigh fading, it is assumed that the (unvectorized) channel $\mathbf{H} \in \mathbb{C}^{2^k \times 2^k}$ (cf. (50), (52)) has elements that are independent standard complex Gaussian random variables, e.g., $[\mathbf{H}]_{i,j} \sim \mathcal{N}(0, 1)$ IID. We consider the channel model in (52) under the Rayleigh fading assumption and additionally assume the noise is AWGN, e.g., $\mathbf{n} \sim \mathcal{N}(0, \sigma^2\mathbf{I}_{TM})$. In the remainder of this section, we derive a noncoherent detection rule under these assumptions. In the sequel, we evaluate the code's performance under the same model.

There are several derivations of the ML detection rule for noncoherent space-time block codes in the Rayleigh fading/Gaussian noise setting [18], [36]. We will make the general assumption that symbols are transmitted according to a uniform distribution over the constellation $\mathcal{C}$ independently over channel realizations. In this setting, ML detection corresponds with maximum a posteriori (MAP) detection. The general maximum likelihood (ML) rule for noncoherent MIMO detection in Rayleigh fading is known to be a quadratic receiver depending on the channel and noise covariance matrices [36]. In this section, we use the properties of stabilizer codes to provide an alternative derivation that yields a receiver that resembles the quantum code's recovery operation.

It was shown in (Section IV-B) that the vectorized channel matrix $\overline{\mathbf{H}}$ could be written as a stochastic combination of the 4^k basis matrices $E^k \subset \mathcal{P}_{2k+1}$ defined in (60) and enumerated in (73). We now derive specific formulae for the expansion of $\overline{\mathbf{H}}$ explicitly. Recall that $M = 2^k$. For the remainder of this section, fix $k \in \mathbb{N}$ and let $E \triangleq E^k$ and $\mathbf{E}_i \triangleq \mathbf{E}_i^k$ (drop the superscripts). The matrices E are unitary, Hermitian, and form a complete, orthogonal basis for the subspace of matrices of the form $\overline{\mathbf{H}} = \mathbf{I}_{2^{k+1}} \otimes \mathbf{R}$, where $\mathbf{R} \in \mathbb{C}^{2^k \times 2^k}$ is arbitrary. The E are not normalized; note that for any i, we have $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i) = 2^{2k+1}$, or, equivalently $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i) = MT$. Thus

$$\overline{\mathbf{H}} = \sum_{i=0}^{4^k-1} \frac{\mathrm{Tr}\left(\mathbf{E}_i^{\mathrm{H}}\overline{\mathbf{H}}\right)\mathbf{E}_i}{MT}. \tag{81}$$

The scalar coefficients $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}}\overline{\mathbf{H}})$ are random variables. Applying the definition of $\mathbf{E}_i$ in (73) and the definition $\overline{\mathbf{H}}$ in (51) gives $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}}\overline{\mathbf{H}}) = \mathrm{Tr}((\mathbf{I}_T \otimes \mathbf{\Gamma}_k(i)^{\mathrm{H}})(\mathbf{I}_T \otimes \mathbf{H}))$, where $\mathbf{\Gamma}_k(i)$ was defined in (72). Let

$$c_i = \mathrm{Tr}\left(\mathbf{\Gamma}_k(i)^{\mathrm{H}}\mathbf{H}\right)/M, \tag{82}$$

and $\mathbf{h} = \mathrm{vec}(\mathbf{H})$. Note that by vectorization this is equivalent to $c_i = \mathrm{vec}(\mathbf{\Gamma}_k(i))^{\mathrm{H}}\mathbf{h}/M$. The bilinearity and trace properties of the Kronecker product can be used to show that $\mathrm{Tr}(\mathbf{E}_i^{\mathrm{H}}\overline{\mathbf{H}}) = (MT)c_i$. Thus (81) is equivalent to

$$\overline{\mathbf{H}} = \sum_{i=0}^{4^k-1} c_i\mathbf{E}_i. \tag{83}$$

The c_i are random variables. In fact, they are the (scalar) linear projections of a joint Gaussian random variable $\mathbf{H}$ onto an orthogonal basis. Thus the c_i are jointly Gaussian. By linearity of expectation and trace, we have $\mathbb{E}[c_i] = 0$. Consider

$$\begin{aligned}
\mathbb{E}\left[c_ic_j^{\mathrm{H}}\right] &= \frac{1}{M^2}\mathbb{E}\left[\mathrm{vec}(\mathbf{\Gamma}_k(i))^{\mathrm{H}}\mathbf{h}\mathbf{h}^{\mathrm{H}}\,\mathrm{vec}(\mathbf{\Gamma}_k(j))\right] &(84)\\
&= \frac{1}{M^2}\,\mathrm{vec}(\mathbf{\Gamma}_k(i))^{\mathrm{H}}\mathbb{E}\left[\mathbf{h}\mathbf{h}^{\mathrm{H}}\right]\mathrm{vec}(\mathbf{\Gamma}_k(j)) &(85)\\
&= \frac{1}{M^2}\,\mathrm{vec}(\mathbf{\Gamma}_k(i))^{\mathrm{H}}\mathbf{I}_{2^k}\,\mathrm{vec}(\mathbf{\Gamma}_k(j)) &(86)\\
&= \frac{1}{M^2}\,\mathrm{Tr}\left(\mathbf{\Gamma}_k(i)^{\mathrm{H}}\mathbf{\Gamma}_k(j)\right) &(87)\\
&= \frac{\mathbb{1}_{i=j}}{M}, &(88)
\end{aligned}$$

where (84) follows from the definition of c_i, (85) follows from the linearity of expectation, (86) follows from the Rayleigh fading assumption on $\mathbf{H}$, (87) follows from identifying the trace inner product as equivalent to the standard inner product between vectorizations, and finally, (88) follows from the fact that if $i \neq j$, the orthogonality of Pauli group elements gives $\mathrm{Tr}(\mathbf{\Gamma}_k(i)^{\mathrm{H}}\mathbf{\Gamma}_k(j)) = \mathbf{0}$. Meanwhile the Pauli group elements are unitary so $\mathbf{\Gamma}_k(i)^{\mathrm{H}}\mathbf{\Gamma}_k(i) = \mathbf{I}_{2^k}$. The vector of scalar projections thus has $\mathbf{c} = [c_0, c_1, \ldots, c_{4^k-1}]^{\mathrm{T}}$ and has $\mathbf{c} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{4^k}/M)$.

At finite SNR, substituting $\mathbf{t} = \sqrt{T}\mathbf{C}\mathbf{s}$ into the vectorized channel model (52) gives $\mathbf{v} = \sqrt{T}\overline{\mathbf{H}}\mathbf{C}\mathbf{s} + \mathbf{n}$, where $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2\mathbf{I}_{2^{2k+1}})$. Note that since $\mathbf{C}$ is a semi-unitary matrix spanning the subspace defined by the projector $\mathbf{P}_0$, we have $\mathbf{P}_0\mathbf{C} = \mathbf{C}$ and thus

$$\mathbf{v} = \sqrt{T}\overline{\mathbf{H}}\mathbf{P}_0\mathbf{C}\mathbf{s} + \mathbf{n}. \tag{89}$$

Using (83) we have

$$\mathbf{v} = \sum_{i=0}^{M^2-1} c_i\sqrt{T}\mathbf{E}_i\mathbf{P}_0\mathbf{C}\mathbf{s} + \mathbf{n}. \tag{90}$$

Let $f_{\mathbf{v}|\mathbf{s}}$ denote the conditional probability density function of the received signal $\mathbf{v}$ given the input symbol $\mathbf{s}$. We seek to solve the maximum likelihood detection problem

$$\hat{\mathbf{s}}_{\mathrm{ML}} \overset{\Delta}{=} \operatorname*{arg\,max}_{\mathbf{s}\in\mathcal{C}} f_{\mathbf{v}|\mathbf{s}}(\mathbf{v}|\mathbf{s}). \tag{91}$$

By assumption, the channel $\mathbf{H}$ is independent of the noise, and thus $\mathbf{n} \perp\!\!\!\perp \mathbf{c}$. Recall that the set of projectors $\{\mathbf{P}_0, \ldots, \mathbf{P}_{M^2-1}\}$ defined in (76) are orthogonal and complete. Let $\hat{\mathbf{q}}_i$ denote the vector projection of the received signal $\mathbf{v}$ onto the subspace spanned by the projector $\mathbf{P}_i$ via

$$\hat{\mathbf{q}}_i = \mathbf{P}_i\mathbf{v}. \tag{92}$$

Via the orthogonality and completeness of the projectors, the set of vectors $\hat{\mathbf{q}}_0, \hat{\mathbf{q}}_1, \ldots, \hat{\mathbf{q}}_{M^2-1}$ is a sufficient statistic of $\mathbf{v}$. Recall the fact that the syndrome of $\mathbf{E}_i$ is the binary expansion of i (cf. Theorem 5). This implies, via the identities (23) and (24) and the orthogonality of the projectors (76), that

$$\mathbf{P}_z\mathbf{E}_i\mathbf{P}_0 = \begin{cases} \mathbf{E}_i\mathbf{P}_0, & \text{if } i = z \\ \mathbf{0}_{2^{2k+1}}, & \text{if } i \neq z. \end{cases} \tag{93}$$

Substituting the received signal model (90) for $\mathbf{v}$ in (92), recognizing that $\mathbf{P}_0\mathbf{C} = \mathbf{C}$, and applying (93) gives

$$\hat{\mathbf{q}}_i = c_i\sqrt{T}\mathbf{E}_i\mathbf{C}\mathbf{s} + \hat{\mathbf{n}}_i \tag{94}$$

where $\hat{\mathbf{n}}_i \sim \mathcal{N}_{\mathrm{C}}(\mathbf{0}, \sigma^2\mathbf{P}_i)$ is the projected noise. The orthogonality of the projectors guarantees that the $\hat{\mathbf{n}}_i$ are mutually independent. In direct analogy to correction operation in stabilizer codes correction, premultiplying $\hat{\mathbf{q}}_i$ by $\mathbf{E}_i^{\mathrm{H}}$ rotates $\hat{\mathbf{q}}_i$ back into the codespace. Let $\tilde{\mathbf{q}}_i = \mathbf{E}_i^{\mathrm{H}}\hat{\mathbf{q}}_i$, and likewise let $\tilde{\mathbf{n}}_i = \mathbf{E}_i^{\mathrm{H}}\hat{\mathbf{n}}_i$. Since the $\mathbf{E}_i$ are unitary, $\tilde{\mathbf{q}}_0, \tilde{\mathbf{q}}_1, \ldots, \tilde{\mathbf{q}}_{M^2-1}$ are a sufficient statistic for $\mathbf{v}$. Using (94), the result of the rotation is

$$\tilde{\mathbf{q}}_i = c_i\sqrt{T}\mathbf{C}\mathbf{s} + \tilde{\mathbf{n}}_i. \tag{95}$$

By direct substitution and the linearity of expectation, we have $\mathbb{E}[\tilde{\mathbf{n}}_i] = \mathbf{0}$. Since $\tilde{\mathbf{n}}_i = \mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\mathbf{n}$ and $\mathbb{E}[\mathbf{n}\mathbf{n}^{\mathrm{H}}] = \mathbf{I}_{MT}\sigma^2$, we have $\mathbb{E}[\tilde{\mathbf{n}}_i\tilde{\mathbf{n}}_i^{\mathrm{H}}] = \mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\mathbf{E}_i$. For all i we have

$$\begin{aligned} \mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\mathbf{E}_i &= \mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\mathbf{P}_i\mathbf{E}_i && (96)\\ &= \big(\mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\big)\big(\mathbf{E}_i^{\mathrm{H}}\mathbf{P}_i\big)^{\mathrm{H}} && (97)\\ &= \big(\mathbf{P}_0\mathbf{E}_i^{\mathrm{H}}\big)\big(\mathbf{P}_0\mathbf{E}_i^{\mathrm{H}}\big)^{\mathrm{H}} && (98)\\ &= \mathbf{P}_0\mathbf{E}_i^{\mathrm{H}}\mathbf{E}_i\mathbf{P}_0 && (99)\\ &= \mathbf{P}_0, && (100) \end{aligned}$$

where (96) follows from the idempotence of projectors, (98) follows from the conjugate of (23), and finally (100) follows from the fact that the Pauli group elements are unitary and the idempotence of projectors. Since the $\hat{\mathbf{n}}_i$ are mutually independent, we have the noise terms $\tilde{\mathbf{n}}_i \sim \mathcal{N}_{\mathrm{C}}(\mathbf{0}, \sigma^2\mathbf{P}_0)$ IID. To solve the detection problem (91), it is sufficient (cf. [41, Chapter 2]) to consider the scalar projections of the $\tilde{\mathbf{q}}_i$ onto the column space of the semi-unitary $\mathbf{C}$. Define $\mathbf{q}_i = \mathbf{C}^{\mathrm{H}}\tilde{\mathbf{q}}_i/\sqrt{2}$, let $\mathbf{w}_i = \mathbf{C}^{\mathrm{H}}\tilde{\mathbf{n}}_i/\sqrt{2}$, and let $\hat{\mathbf{c}} = \sqrt{M}\mathbf{c} \sim \mathcal{N}_{\mathrm{C}}(\mathbf{0}, \mathbf{I}_{4^k})$. Since $\mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}_2$, the definition of $\tilde{\mathbf{q}}_i$ in (95) gives

$$\mathbf{q}_i = \hat{c}_i\mathbf{s} + \mathbf{w}_i, \tag{101}$$

where we used the fact that $T = 2M$. Note that $\mathbb{E}[\mathbf{w}_i] = 0$ and that since $\mathbf{C}^{\mathrm{H}}\mathbf{P}_0\mathbf{C} = \mathbf{I}_2$ (cf. (77)) we have $\mathbb{E}[\mathbf{w}_i\mathbf{w}_i^{\mathrm{H}}] = \sigma^2\mathbf{I}_2/2$. Since the $\tilde{\mathbf{n}}_i$ are mutually independent, $\mathbf{w}_i \sim \mathcal{N}_{\mathrm{C}}(\mathbf{0}, \frac{\sigma^2}{2}\mathbf{I}_2)$ IID.

Define $\mathbf{w} = [\mathbf{w}_1^{\mathrm{T}}, \mathbf{w}_2^{\mathrm{T}}, \ldots, \mathbf{w}_{4^k-1}^{\mathrm{T}}]^{\mathrm{T}}$, and let $\mathbf{z} = [\hat{\mathbf{c}}^{\mathrm{T}}, \mathbf{w}^{\mathrm{T}}]^{\mathrm{T}}$. The ML rule will follow from writing the sufficient statistic $\mathbf{q} = [\mathbf{q}_0^{\mathrm{T}}, \mathbf{q}_1^{\mathrm{T}}, \ldots, \mathbf{q}_{4^k-1}^{\mathrm{T}}]^{\mathrm{T}}$ as a linear function of the joint Gaussian random vector $\mathbf{z}$. The definitions of $\hat{\mathbf{c}}$, $\mathbf{w}$, and the aforementioned independence relationships guarantee that

$$\mathbf{z} \sim \mathcal{N}_{\mathrm{C}}(\mathbf{0}, \mathbf{R}) \tag{102a}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{I}_{2^{2k}} & \mathbf{0}_{2^{2k}\times 2^{2k+1}} \\ \mathbf{0}_{2^{2k+1}\times 2^{2k}} & \frac{\sigma^2}{2}\mathbf{I}_{2^{2k+1}} \end{bmatrix}. \tag{102b}$$

Define the matrix $\mathbf{M} \in \mathbb{C}^{2^{2k+1}\times(3)4^k}$ via

$$\mathbf{M} = \big[\big(\mathbf{I}_{4^k} \otimes \mathbf{s}\big)\ \mathbf{I}_{2^{2k+1}}\big]. \tag{103}$$

Using (101) and the definition of $\mathbf{z}$, it follows that

$$\mathbf{q} = \mathbf{M}\mathbf{z}. \tag{104}$$

Thus for a given transmit symbol $\mathbf{s}$, $\mathbf{q}$ is a joint Gaussian random vector with $\mathbb{E}[\mathbf{q}] = \mathbf{0}$. Defining $\mathbb{E}[\mathbf{q}\mathbf{q}^{\mathrm{H}}] = \mathbf{Q}$ we have $\mathbf{Q} = \mathbf{M}\mathbf{R}\mathbf{M}^{\mathrm{H}}$, or, equivalently

$$\mathbf{Q} = \mathbf{I}_{4^k} \otimes \Big[\big(\mathbf{s}\mathbf{s}^{\mathrm{H}}\big) + \tfrac{\sigma^2}{2}\mathbf{I}_2\Big]. \tag{105}$$

By (105),

$$\mathbf{q} \sim \mathcal{N}_{\mathrm{C}}\big(\mathbf{0}_{2^{2k+1}}, \mathbf{Q}\big). \tag{106}$$

Since $\mathbf{q}$ is a sufficient statistic to estimate the transmitted symbol $\mathbf{s}$, the decoding rule (91) is equivalent to

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \operatorname*{arg\,max}_{\mathbf{s}\in\mathcal{C}} f_{\mathbf{q}|\mathbf{s}}(\mathbf{q}|\mathbf{s}) \tag{107a}$$

where

$$f_{\mathbf{q}|\mathbf{s}}(\mathbf{q}|\mathbf{s}) = \frac{\exp\big(-\mathbf{q}^{\mathrm{H}}\mathbf{Q}^{-1}\mathbf{q}\big)}{\pi^{2^{2k+1}}\det(\mathbf{Q})}. \tag{107b}$$

The properties of determinants of the Kronecker products, together with the assumption that $\|\mathbf{s}\|_2 = 1$ for all $\mathbf{s} \in \mathcal{C}$, can be used to show that the denominator of (107b) does not depend on the transmitted symbol $\mathbf{s}$.

The decision rule reduces to

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \operatorname*{arg\,min}_{\mathbf{s}\in\mathcal{C}} \mathbf{q}^{\mathrm{H}}\mathbf{Q}^{-1}\mathbf{q}. \tag{108}$$

Note that $\mathbf{Q}^{-1} = \mathbf{I}_{4^k} \otimes (\mathbf{s}\mathbf{s}^{\mathrm{H}} + \frac{\sigma^2}{2}\mathbf{I}_2)^{-1}$. Define $\mathbf{U}_{\mathbf{s}} = (\mathbf{s}\mathbf{s}^{\mathrm{H}} + \frac{\sigma^2}{2}\mathbf{I}_2)^{-1}$. Carrying out the matrix inversion explicitly and using $\|\mathbf{s}\|_2 = 1$ gives

$$\mathbf{U}_{\mathbf{s}} = \frac{1}{\frac{\sigma^2}{2}\left(\frac{\sigma^2}{2} + 1\right)}\left(\frac{\sigma^2}{2}\mathbf{I}_2 + \left(\mathbf{I}_2 - \mathbf{s}\mathbf{s}^{\mathrm{H}}\right)\right). \tag{109}$$

Substituting the expression for $\mathbf{Q}^{-1}$ and the definition of $\mathbf{U}_{\mathbf{s}}$ into (108) gives $\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\min_{\mathbf{s}\in\mathcal{C}} \mathbf{q}^{\mathrm{H}}(\mathbf{I}_{4^k} \otimes \mathbf{U}_{\mathbf{s}})\mathbf{q}$. Using (109), the detection rule becomes

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}\in\mathcal{C}} \mathbf{q}^{\mathrm{H}}\left(\mathbf{I}_{4^k} \otimes \mathbf{s}\mathbf{s}^{\mathrm{H}}\right)\mathbf{q}. \tag{110}$$

As anticipated (cf. [36]), the decision rule reduces to a quadratic form. Substituting the definition of $\mathbf{q}$ into (110) yields

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}\in\mathcal{C}} \mathbf{s}^{\mathrm{H}}\left(\sum_{n=0}^{4^k-1} \mathbf{q}_k\mathbf{q}_k^{\mathrm{H}}\right)\mathbf{s}. \tag{111}$$

Making the substitution $\mathbf{q}_i = \mathbf{C}^{\mathrm{H}}\mathbf{E}_i\mathbf{P}_i\mathbf{v}/\sqrt{2}$ demonstrates that (111) is equivalent to

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}\in\mathcal{C}} \mathbf{s}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right)\mathbf{C}\mathbf{s}, \tag{112}$$

where the emulated recovery operation $f_{\mathcal{R}_C}$ is as given in (79). This demonstrates that the optimal detection rule for Rayleigh fading is precisely the emulated recovery operation (47). Thus, the detection rule based on the emulated recovery operation is guaranteed to perform perfectly at infinite SNR for all channel distributions with $\mathbb{P}[\mathbf{H} = \mathbf{0}] = 0$. Meanwhile, the rule actually corresponds to the ML detector for Rayleigh fading and finite SNR. Computing the ML rule exactly has a complexity of $\mathcal{O}(M^4 + M^2|\mathcal{C}|)$ complex multiply-adds. Naively applying the generalized likelihood ratio test in [36] gives a complexity that scales like $\mathcal{O}(M^4|\mathcal{C}|)$. Thus, at higher rates, the quantum-inspired receiver is advantageous from a complexity perspective. In the next section, we proceed with a performance analysis and numerical results.

VII. ANALYSIS AND NUMERICAL RESULTS

In this section, we analyze the space-time codes designed via the techniques in Section V and compare their performance with coherent and noncoherent approaches with comparable rates and latency constraints. Explicitly, we consider schemes that use an (approximately) equal number of bits per coherence interval and consider schemes that code over only a single coherence interval.

A. ANALYSIS

It turns out that with maximum likelihood detection this code achieves full diversity; at high SNR the error rate decreases at a rate of M^2 decades per 10dB increase in SNR. This is exhibited by a standard calculation of a bound on the pairwise error probability (cf. Appendix E). Define the pairwise error probability $\mathbb{P}[\mathbf{s}_p \to \mathbf{s}_q]$ as the probability the decoding rule (111) detects $\mathbf{s}_q \in \mathcal{C}$ when the transmitted symbol was $\mathbf{s}_p \in \mathcal{C}$, let $\delta_{p,q} = \mathbf{s}_p^{\mathrm{H}}\mathbf{s}_q$, and let $\xi = \mathrm{SNR}^{-1}$. The pairwise error probability can be shown to be bounded via

$$\mathbb{P}\left[\mathbf{s}_p \to \mathbf{s}_q\right] \leq \left(1 - \frac{1 - |\delta_{p,q}|^2}{\xi^2 + 2\xi + \left(1 - |\delta_{p,q}|^2\right)}\right)^{M^2}. \tag{113}$$

Expanding around $\xi = 0$ (i.e., infinite SNR) demonstrates that a diversity of M^2 is asymptotically achieved. As shown in Appendix E, the bound (113) follows from the fact that decoding rule (111) is the sum of M^2 independent, identically distributed random variables. As shown in Section VI, this followed exactly from the fact that each error operator had a unique syndrome with respect to the stabilizer generators.

The bound also motivates a particular choice of symbol constellation, $\mathcal{C}$. The right hand side of (113) is an increasing function of $|\delta_{p,q}|^2$. We choose the input constellation $\mathcal{C}$ to minimize the maximum $|\delta_{p,q}|^2$. Formally, to encode $\log_2(N)$ bits per codeword, we choose $\mathcal{C}$ as a Grassmannian line packing via the subset selection program

$$\mathcal{C} = \underset{\substack{\hat{\mathcal{C}}\subset\left\{\mathbf{s}_i\in\mathbb{C}^2 \mid \|\mathbf{s}_i\|_2=1\right\} \\ |\hat{\mathcal{C}}|=N}}{\arg\min} \; \max_{\substack{\mathbf{s}_x,\mathbf{s}_z\in\hat{\mathcal{C}} \\ x\neq z}} |\mathbf{s}_z^{\mathrm{H}}\mathbf{s}_x|^2, \tag{114}$$

which minimizes the maximum pairwise error probability and thus minimizes an upper bound on the probability of error. We reiterate that the constellation packing is a packing of one-dimensional subspaces in a low-dimensional space ($\mathbf{s} \in \mathbb{C}^2$), in contrast to the many codes in the literature based on packing the rowspace of the matrix codewords $\mathbf{T} \in \mathbb{C}^{M\times T}$ (cf. Section III-A). The difficulty of computing the bound (113) is essentially just the difficulty of computing $|\delta_{p,q}|$. This complexity depends only on the rate of the code and is independent of the number of antennas.

B. PERFORMANCE EVALUATION

In our simulations, we considered systems with $M = \{2, 4, 8\}$ antennas at both the transmitter and the receiver. Each qubit symbol constellation is in principle arbitrary and generates a different code. We used Grassmannian packings of sizes 4 and 8 in $\mathbb{C}^2$ for the symbol constellations, giving two codes (at two different rates) for each choice of M. The packings we used are listed in [42]. To gauge the quality of these packings, it is useful to compare the packing's maximum $|\delta_{m,n}|$ to lower bounds found in the literature [43]. The chordal distance bound in [44] (equivalently the spectral distance bound in [43] for one-dimensional packings) can be used to show that, for a packing of N one-dimensional subspaces in a two-dimensional complex vector space

$$\max_{m\neq n} |\delta_{m,n}|^2 \geq \left(1 - \frac{N}{2(N-1)}\right). \tag{115}$$

The 4 packing in [42] has $|\delta_{m,n}|$ (numerically) constant for all $m \neq n$, achieves this bound (to nearly machine precision), and is thus optimal. This bound is (provably) not achievable

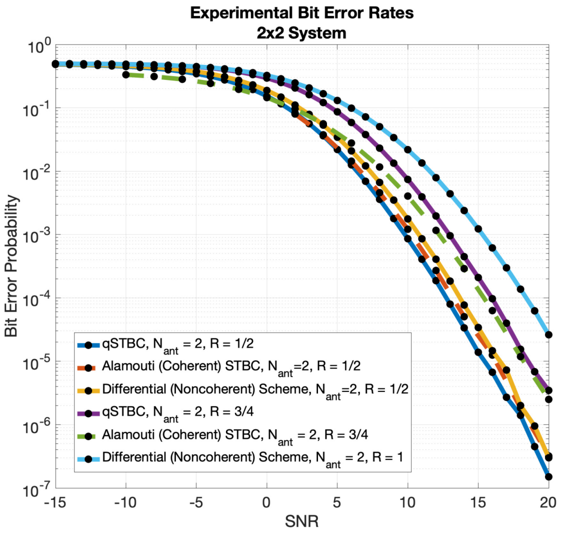

FIGURE 3. Bit error rates for the 2x2 system at two different rates. We simulated 10 million channel realizations. The quantum-inspired space-time block code is labeled qSTBC. At rate $r = 1/2$, the quantum-inspired code outperforms both the Alamouti and the differential scheme. The rate $r = 3/4$ code significantly outperforms the $r = 1$ differential code, however both are outperformed by the Alamouti approach at $r = 3/4$.

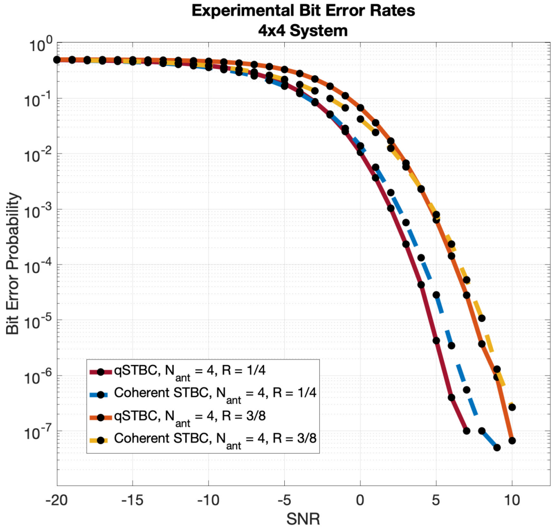

FIGURE 4. Bit error rates for the 4x4 system at two different rates. We simulated 10 million channel realizations. The quantum-inspired space-time block code is labeled qSTBC.

for the 8 packing [43], [44]. The maximum $|\delta_{m,n}|^2$ is a factor of about 1.65 times the bound. We did not conduct an extensive literature search for the best known 8 packing. We also did not attempt to develop our own packings via numerical optimization (cf. [43] and references).

For systems of 2, 4, and 8 antennas, we compare our performance to coherent and noncoherent space-time code constructions with comparable bitrates that can be decoded after one coherence interval T. For coherent approaches, we assume that an equal amount of power is used for training and communication, with constant noise (and thus SNR) throughout the coherence interval. We found that for $M + 2$, our lower rate code (derived from the 4 qubit symbol input constellation) outperforms

1) M = 2

Bit error rate results for the $M = 2$, $T = 4$ are shown in Fig. 3. We compared the stabilizer-based, non-coherent construction with a coherent scheme based on the Alamouti code (for 2×2 systems) at spectral efficiency rates of $r = 1/2$ and $r = 3/4$ bits/channel use [45]. In the first two channel uses, the transmitter transmits training symbols $[1, 1]^{\mathrm{T}}/\sqrt{2}$ and $[1, -1]^{\mathrm{T}}/\sqrt{2}$ and the receiver forms an estimate of $\mathbf{H}$. We then use the Alamouti scheme to transmit one space-time symbol $\mathbf{s} \in \mathbb{C}^2$ over the remaining two channel uses in the coherence interval. We encode in $\mathbf{s}$ two binary phase-shift keying (BPSK) symbols for the rate $r = 1/2$ approach. For the rate $r = 3/4$ approach, we encoded one quadrature phase-shift keying (QPSK) symbol and one BPSK symbol. Similarly, we compare to an approach using differential unitary group codes, as outlined in [28]. At both $r = 1/2$ and $r = 1$, the first two channel uses are used for the 2×2 reference matrix, and no information is transmitted. With the next two channel uses we transmit a single differentially encoded 2×2 matrix drawn from an appropriately sized constellation. For $r = 1/2$, this constellation is a group code over the QPSK constellation. For $r = 1$, this constellation is a dicylic group code generated over 16-PSK.

We found that the lower rate code (derived from the 4 qubit symbol input constellation) outperformed both the Alamouti-based coherent scheme as well as the differential approach. The higher rate code outperforms the differential scheme, but seems to perform worse than the Alamouti approach at all SNRs.

2) M = 4

Bit error rate results for the $M = 4$, $T = 8$ are shown in Fig. 4. We compared the stabilizer-based, non-coherent construction with a coherent scheme at spectral efficiency rates of $r = 1/4$ and $r = 3/8$ bits/channel use [46]. For the coherent approaches, the transmitters are assumed to transmit unit-power, orthogonal training symbols during the first four time instants. The following four instants are used for communication. We specifically use the 4×4 codewords given in (4) in [46]. The receiver assumes that the channel estimate is the true channel and decodes using minimum distance decoding (i.e., it finds $\arg \min_{\mathbf{T}} \|\mathbf{H}_{\text{est}}\mathbf{T} - \mathbf{V}\|_F$). The codewords in [46] accommodate up to four transmit symbols. For the $r = 1/8$ system, we encode two BPSK symbols, repeated one time each. For the $r = 3/8$ system, we encoded three BPSK symbols, one of which was repeated once.

When $M = 4$, both the lower and higher rate codes outperform the coherent approach at higher SNRs. At moderate to low SNRs, the coherent approaches perform better.

3) M = 8

Bit error rate results for the $M = 8$, $T = 16$ are shown in Fig. 5. We compared the stabilizer-based, non-coherent

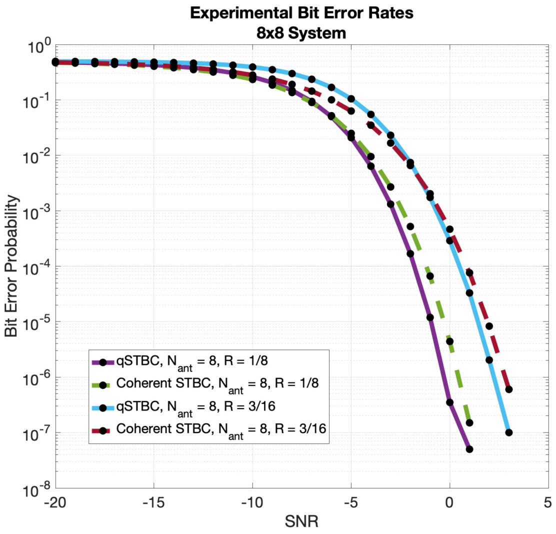

FIGURE 5. Bit error rates for the 8x8 system at two different rates. We simulated 10 million channel realizations. The quantum-inspired space-time block code is labeled qSTBC.

construction with a coherent scheme at spectral efficiency rates of $r = 1/8$ and $r = 3/16$ bits/channel use [47]. For the coherent approaches the transmitters are assumed to transmit unit-power training symbols during the first four time instants. The following four instants are used for communication. The receiver assumes that the channel estimate is the true channel and decodes using minimum distance decoding (i.e., it finds $\arg\min_{\mathbf{T}} \|\mathbf{H}_{\text{est}}\mathbf{T} - \mathbf{V}\|_F$). The 8×8 codewords in [47] accommodate up to four transmitted symbols. For the $r = 1/8$ system, we encode two BPSK symbols, repeated one time each. For the $r = 3/16$ system, we encoded 3 BPSK symbols, one of which was repeated once. When $M = 8$, both the lower and higher rate codes outperform the coherent approach at higher SNRs. At moderate to low SNRs, the coherent approaches perform better. Generally speaking, the amount of improvement of the quantum space-time block code over the coherent approach at high SNR seems to improve as the rates decrease for a fixed number of antennas.

VIII. CONCLUSION

We first summarize our contributions and results before discussing opportunities for future work.

A. CONTRIBUTIONS AND RESULTS

In this paper, we developed a framework to use quantum error correcting codes in classical noncoherent communication. We demonstrated that a rich class of classical channels (linear channels with additive noise) are amenable to a characterization as an "emulated quantum channel." We proposed a novel modulation scheme where data was encoded in the mathematical structure of quantum error correcting codes. We extended the quantum error correction conditions to the classical setting. We then demonstrated that when a certain quantum code design problem could be solved, the aforementioned modulation scheme coupled with a quantum-inspired detection rule, dubbed the "emulated recovery operation", could achieve zero-error noncoherent communication. Our approach, and these performance guarantees, are applicable to a rich class of noncoherent communication problems over linear channels.

We then applied the new coding framework to noncoherent MIMO wireless communication. We solved the relevant quantum code design problem for this channel at infinite SNR using the stabilizer formalism. This led to a novel family of space-time block codes. We used the properties of stabilizer codes to derive the maximum likelihood detection rule for the Rayleigh fading case at finite SNR. The maximum likelihood detector was shown to correspond to the emulated recovery operation, and was thus guaranteed to perform perfectly at an infinite SNR, regardless of the distribution of the channel matrix. Furthermore, the code design led to a maximum likelihood decoder with reduced computational complexity. We used the derived maximum likelihood decoder to derive a bound on the probability of error. This bound is relatively easy to compute and analyze, even for systems with large numbers of antennas. Both the simplified decoder and the easy-to-compute bound arise from the fact that our construction maps low-dimensional Grassmann packings (the constellation) to packings in a higher-dimensional space (the space-time codewords).

In the Rayleigh/Gaussian setting, our numerical results demonstrated that, for the rates and antenna configurations we evaluated, the family of codes achieves good performance with respect to comparable coherent and noncoherent schemes. While our new codes tended to underperform their exemplary coherent and noncoherent counterparts at low SNRs, they achieved modest gains at higher SNRs. It is difficult to derive an exact analytical characterization of the error probabilities of the various coherent and noncoherent signaling schemes we evaluated. Generally speaking, analytical insight into the performance of these schemes relies on upper bounds (cf. [36]). For all rates and antenna geometries we evaluated, we found that our construction lead to a unitary space-time block code. We conjecture that this holds generally, but do not yet have a proof. The differential schemes we evaluated can be shown to be equivalent to a unitary space-time block code. Generally speaking, the performance of these codes at high SNR is dominated by their so-called *product diversity* (cf. [25]), which is related to the pairwise separation of transmit space-time symbols in signal space. For the rates and geometries we evaluated, our construction might lead to a higher product diversity, in particular relative to the differential scheme we evaluated.

B. FUTURE WORK

Our results leave several open questions, including an exact characterization of the relationship between the error probabilities of the various space-time block coding schemes we evaluated. Additionally, using our framework, quantum error correction may be applied to other classical communication problems. In particular, for the noiseless noncoherent MIMO channel, we were able to extend our stabilizer code design to the case of $M = T = 2$, a regime where coherent communication is impossible. The code was a slight perturbation of the one presented here and achieved finite SNR performance equivalent to the $M = 2$, $T = 4$ scheme of Section V-A. In particular, the quantum code used was degenerate (cf. Section III-B).

In applying our techniques to practical classical communication problems, we do not expect to be able to design a quantum code that can correct all possible errors in the emulated quantum channel. In particular, our work demonstrates that this is impossible when the additive noise is not deterministically zero or when the original channel has a non-infinite capacity. If one attempts to use QEC in such a setting, we suggest designing a code for a subset of the error operators encountered, and then attempting to design the optimal decoder. Often times, designing the optimal decoder is intractable, or the decoder ends up being too computationally complex. In this case, it may be effective to use a suboptimal detection rule inspired by the emulated recovery operation.

Some theoretical insight from Section III may be applied to both quantum and classical information theory. In particular, several results can be drawn proving that some quantum codes do not exist. Assume "real" quantum channel's Kraus operators can be shown to correspond to some classical channel's emulated quantum channel. If the corresponding classical channel is not infinite capacity, or has additive noise, our work demonstrates that it is impossible to design a quantum error correcting code for the quantum channel. Similarly, with our approach, one could apply "achievability" results from a quantum channel to a classical channel whose emulated Kraus operators correspond to those of the quantum channel. These information/coding theoretic implications are an opportunity for future work.

APPENDIX

A. PROOF OF THEOREM 1 AND COROLLARY 1

We first state a lemma.

Lemma 1: Assume $\mathbf{t} \in \mathbb{C}^m$ is an arbitrary vector with $\mathbf{t}^{\mathrm{H}}\mathbf{t} = 1$. Let $\mathbf{e}_i$ denote the i^{th} standard unit vector in $\mathbb{C}^m$. Let $\mathbf{J}_{i,j} = \mathbf{e}_i\mathbf{e}_j^{\mathrm{H}}$. It can be shown that

$$\mathbf{I}_{m^2} = \sum_{i=0}^{m^2-1}\sum_{j=0}^{m^2-1} \mathbf{J}_{i,j}\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{J}_{i,j}^{\mathrm{H}}. \tag{116}$$

Proof: For every i, the quantity $\sum_{j=0}^{m^2-1} \mathbf{J}_{i,j}\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{J}_{i,j}^{\mathrm{H}} = \mathbf{e}_i\mathbf{e}_i^{\mathrm{H}}$. ■

Corollary 2: If $\mathbf{\Phi} \in \mathbb{C}^{m\times m}$ with $\mathbf{\Phi} \succeq 0$, then

$$\sum_{i=0}^{m^2-1}\sum_{j=0}^{m^2-1} \mathbf{J}_{i,j}\mathbf{\Phi}\mathbf{J}_{i,j}^{\mathrm{H}} = \mathrm{Tr}(\mathbf{\Phi})\mathbf{I}_m. \tag{117}$$

The corollary follows from writing $\mathbf{\Phi}$ as a sum of rank 1 matrices.

1) PROOF OF THEOREM 1

Proof: Let $\overline{\mathbf{H}}_0 = \overline{\mathbf{H}} - \overline{\mathbf{H}}_\mu$. Consider an arbitrary, orthonormal (with respect to the trace inner product) basis for $m \times m$ matrices, denoted $\mathbf{\Theta}_i$, where $i \in \{0, 1, \ldots, m^2 - 1\}$.

Define

$$c_i = \mathrm{Tr}\Big(\mathbf{\Theta}_{\boldsymbol{i}}\overline{\mathbf{H}}_0\Big) \tag{118}$$

and denote the concatenation $\mathbf{c} = [c_0, c_1, \ldots, c_{m^2-1}]^{\mathrm{T}}$. Note that we have (using block matrix notation)

$$\overline{\mathbf{H}}_0 = \big[\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_{m^2-1}\big]\big(\mathbf{c} \otimes \mathbf{I}_{m^2}\big). \tag{119}$$

Since $\mathbb{E}[\overline{\mathbf{H}}_0] = 0$, $\mathbb{E}[\mathbf{c}] = \mathbf{0}$. Let $\mathbb{E}[\mathbf{c}\mathbf{c}^{\mathrm{H}}] = \mathbf{R}$. Since $\mathbf{R} \succeq 0$, for some unitary $\mathbf{U}$ and positive, real, diagonal matrix $\mathbf{\Sigma}$ we have $\mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}}$. Define $\hat{\mathbf{c}} = \mathbf{U}^{\mathrm{H}}\mathbf{c}$. We have $\mathbb{E}[\hat{\mathbf{c}}] = \mathbf{0}$ and $\mathbb{E}[\hat{\mathbf{c}}\hat{\mathbf{c}}^{\mathrm{H}}] = \mathbf{\Sigma}$. Note that (119) can be written

$$\overline{\mathbf{H}}_0 = \big[\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_{m^2-1}\big]\big(\mathbf{U} \otimes \mathbf{I}_{m^2}\big)\big(\hat{\mathbf{c}} \otimes \mathbf{I}_{m^2}\big). \tag{120}$$

Let $\tilde{\Theta} = \big[\mathbf{\Theta}_0, \mathbf{\Theta}_2, \ldots, \mathbf{\Theta}_{m^2-1}\big](\mathbf{U} \otimes \mathbf{I}_{m^2})$. Define $\hat{\Theta}_i = \sum_{j=0}^{m^2} \Theta_j[\mathbf{U}]_{j,i}$. It follows that $\tilde{\Theta} = \Big[\hat{\mathbf{\Theta}}_0, \hat{\mathbf{\Theta}}_1 \ldots, \hat{\mathbf{\Theta}}_{m^2-1}\Big]$. Let $\hat{c}_i = [\hat{\mathbf{c}}]_i$, and let $\sigma_i = [\mathbf{\Sigma}]_{i,i}$. Thus, using (120) and the new definitions, we have that

$$\overline{\mathbf{H}}_0 = \sum_{i=0}^{m^2-1} \hat{c}_i\hat{\mathbf{\Theta}}_i, \tag{121}$$

where $\mathbb{E}[\hat{c}_i\hat{c}_j^{\mathrm{H}}] = 0$ if $i \neq j$ and $\mathbb{E}[\hat{c}_i\hat{c}_i^{\mathrm{H}}] = \sigma_i$.

Substituting $\mathbf{v} = (\overline{\mathbf{H}}_0 + \overline{\mathbf{H}}_\mu)\mathbf{t} + \mathbf{n}$ and using the definition of $\mathbf{n}$ as well as the independence assumptions gives

$$\begin{aligned}\mathbb{E}\Big[\Big(\overline{\mathbf{H}}\mathbf{t} + \mathbf{n}\Big)\Big(\overline{\mathbf{H}}\mathbf{t} + \mathbf{n}\Big)^{\mathrm{H}}|\mathbf{t}\Big]\\ = \mathbb{E}\Big[\overline{\mathbf{H}}\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}^{\mathrm{H}}|\mathbf{t}\Big] + \mathbb{E}\big[\mathbf{n}\mathbf{n}^{\mathrm{H}}\big].\end{aligned} \tag{122}$$

We attack each of the terms in (122) one at a time.

By assumption $\mathbb{E}[\mathbf{n}\mathbf{n}^{\mathrm{H}}] = \mathbf{\Phi}$. Since $\mathbf{\Phi} \succeq 0$, the matrix square root $\mathbf{\Phi}^{\frac{1}{2}}$ is well-defined. Using the notation from Lemma 1, $\mathbf{\Phi}$ can be written

$$\mathbf{\Phi} = \sum_{i=0}^{m^2-1}\sum_{j=0}^{m^2-1} \mathbf{\Phi}^{\frac{1}{2}}\mathbf{J}_{i,j}\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{J}_{i,j}^{\mathrm{H}}\mathbf{\Phi}^{\frac{1}{2}}. \tag{123}$$

Using the definitions of $\overline{\mathbf{H}}_0$ and $\overline{\mathbf{H}}_\mu$,

$$\mathbb{E}\Big[\overline{\mathbf{H}}\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}^{\mathrm{H}}|\mathbf{t}\Big] = \overline{\mathbf{H}}_\mu\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}_\mu^{\mathrm{H}} + \mathbb{E}\Big[\overline{\mathbf{H}}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}_0^{\mathrm{H}}|\mathbf{t}\Big]. \tag{124}$$

The second term on the right hand side can be simplified using the expansion in (121) via

$$\mathbb{E}\Big[\overline{\mathbf{H}}_0\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}_0^{\mathrm{H}}|\mathbf{t}\Big] = \sum_{i=0}^{m^2-1} \sigma_i\hat{\Theta}_i\mathbf{t}\mathbf{t}^{\mathrm{H}}\hat{\Theta}_i^{\mathrm{H}}. \tag{125}$$

Substitute (125) into (124) to demonstrate that $\mathbb{E}[\overline{\mathbf{H}}\mathbf{t}\mathbf{t}^{\mathrm{H}}\overline{\mathbf{H}}^{\mathrm{H}}|\mathbf{t}]$ may be written as an operator sum. Substituting the resulting expression for (123) and (124) into (122) gives the result, where $\mathbf{E}_i$ is an arbitrary indexing of the operators $\mathbf{H}_\mu$, $\sqrt{\sigma}_i\hat{\Theta}_i$, and $\mathbf{\Phi}^{\frac{1}{2}}\mathbf{J}_{i,j}$. ■

2) PROOF OF COROLLARY 1

Proof: Note first that $\mathrm{Tr}(\mathbf{\Theta}_i^{\mathrm{H}}\overline{\mathbf{H}}) \overset{\mathrm{a.s}}{=} 0$ if and only if both $\mathrm{Tr}(\mathbf{\Theta}_i^{\mathrm{H}}\overline{\mathbf{H}}_0) \overset{\mathrm{a.s}}{=} 0$ and $\mathrm{Tr}(\mathbf{\Theta}_i^{\mathrm{H}}\overline{\mathbf{H}}_\mu) = 0$. The completeness of the basis implies that $\overline{\mathbf{H}}_\mu$ can be expressed as a linear combination of $\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots \mathbf{\Theta}_{z-1}$. Define c_i as in (118). Note that by assumption, $c_i \overset{\mathrm{a.s}}{=} 0$ for all $i \geq z$. Thus

$$\overline{\mathbf{H}}_0 \overset{\mathrm{a.s}}{=} \sum_{i=0}^{z-1} c_i\mathbf{\Theta}_i. \tag{126}$$

Thus, the random matrix $\overline{\mathbf{H}}_0$ can (almost surely) be written as a stochastic linear combination of $\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots \mathbf{\Theta}_{z-1}$. As in the prequel, the moments of the coefficients $\mathbf{c}_i$ can be used to derive the Kraus operators $\{\mathbf{E}_i\}$. Unsurprisingly, these operators can be shown to be in the span of $\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_{z-1}$. To see this, define $\mathbf{c} = [c_0, c_1, \ldots, c_{z-1}]$ (a truncated version of the vector defined in the prequel). Comparing with (119), (126) may be written

$$\overline{\mathbf{H}}_0 \overset{\mathrm{a.s}}{=} \left[\mathbf{\Theta}_0, \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_{z-1}\right]\left(\mathbf{c} \otimes \mathbf{I}_{m^2}\right). \tag{127}$$

Continuing with the argument from (119)) to (122) in this manner (using the lower-dimensional $\mathbf{c} \in \mathbb{C}^z$ and replacing $=$ with $\overset{\mathrm{a.s}}{=}$ as appropriate) proves the corollary. ■

B. PROOF OF THEOREM 3

The next lemma follows directly from Theorem 2.

Lemma 2: If $\mathbb{P}[\mathbf{v} = 0] = 0$, then $f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}}) \overset{\mathrm{a.s}}{\neq} 0$.

Proof: Let the recovery operation $f_{\mathcal{R}_C}$ be as defined in (45). It can be shown that the recovery operation has no nullspace in the sense that if $\mathbf{Q} \succeq 0$, $f_{\mathcal{R}_C}(\mathbf{W}) = \mathbf{0}$ if and only if $\mathbf{W} = \mathbf{0}$. To see the converse, use the definition of $\mathbf{R}_i$ in (41) to show that, for $\mathbf{W} \succeq 0$

$$\mathrm{Tr}\left(f_{\mathcal{R}_C}(\mathbf{W})\right) = \mathrm{Tr}(\mathbf{W}). \tag{128}$$

This implies that $\mathbb{P}[f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}}) = 0] = \mathbb{P}[\mathbf{v}\mathbf{v}^{\mathrm{H}} = 0]$. By assumption $\mathbb{P}[\mathbf{v}\mathbf{v}^{\mathrm{H}} = 0] = 0$, which gives the result. ■

We now prove Theorem 3. *Proof:* Consider the random variable

$$\mathbf{\Lambda} = f_{\mathcal{R}_C}\left(\mathbf{v}\mathbf{v}^{\mathrm{H}}\right). \tag{129}$$

By linearity of conditional expectation

$$\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}] = f_{\mathcal{R}_C}\left(\mathbb{E}\left[\mathbf{v}\mathbf{v}^{\mathrm{H}}|\mathbf{t}\right]\right). \tag{130}$$

Recall that by assumption the conditions of Theorem 2 are satisfied and $\mathbf{t} = \mathbf{P}\mathbf{t}$. Substituting the definition of $f_{\mathcal{R}_C}$ in (45) into (130) and applying (40) gives

$$\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}] = c\mathbf{t}\mathbf{t}^{\mathrm{H}}, \tag{131}$$

for some positive constant c.

Note that by definition $\mathbf{\Lambda} \succeq 0$. Let $\mathbf{t}$ be a vector. Define two projections of PSD matrices via

$$\mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbf{\Lambda}) = \frac{\left(\mathbf{t}^{\mathrm{H}}\mathbf{\Lambda}\mathbf{t}\right)}{\|\mathbf{t}\|_2^4}\mathbf{t}\mathbf{t}^{\mathrm{H}} \tag{132a}$$

and

$$\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda}) = \mathbf{\Lambda} - \mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbf{\Lambda}). \tag{132b}$$

The projection in (132a) projects the row and column spaces of $\mathbf{\Lambda}$ onto the subspace spanned by the vector $\mathbf{t}$, while the projection in (132b) projects $\mathbf{\Lambda}$ onto the orthogonal complement. Since $\mathbf{\Lambda} \succeq 0$ the projections (132a) and (132b) are both PSD. Furthermore

$$\mathbf{\Lambda} = \mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbf{\Lambda}) + \mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda}). \tag{133}$$

Again using linearity of conditional expectation

$$\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}] = \mathbb{E}\left[\mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbf{\Lambda})|\mathbf{t}\right] + \mathbb{E}\left[\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda})|\mathbf{t}\right] \tag{134}$$

By linearity and the fact that $\mathbf{t}\mathbf{t}^{\mathrm{H}}$ is $\sigma(\mathbf{t})$ measurable, we have

$$\mathbb{E}\left[\mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbf{\Lambda})|\mathbf{t}\right] = \mathrm{proj}_{\mathbf{t}\mathbf{t}^{\mathrm{H}}}(\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}]), \tag{135a}$$

$$\mathbb{E}\left[\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda})|\mathbf{t}\right] = \mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}]). \tag{135b}$$

However by (131) $\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}] = c\mathbf{t}\mathbf{t}^{\mathrm{H}}$ and so

$$\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbb{E}[\mathbf{\Lambda}|\mathbf{t}]) = \mathbf{0}, \tag{136}$$

which by (135b) demonstrates that

$$\mathbb{E}\left[\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda})|\mathbf{t}\right] = \mathbf{0}. \tag{137}$$

Since $\mathrm{proj}_{(\mathbf{t}\mathbf{t}^{\mathrm{H}})^\perp}(\mathbf{\Lambda})$ is a random PSD matrix with an expectation equal to zero, it follows that

$$\mathrm{proj}_{\left(\mathbf{t}\mathbf{t}^{\mathrm{H}}\right)^\perp}(\mathbf{\Lambda}) \overset{\mathrm{a.s}}{=} \mathbf{0}. \tag{138}$$

This implies that for some real valued nonnegative random variable γ we have

$$\mathbf{\Lambda} \overset{\mathrm{a.s}}{=} \gamma\mathbf{t}\mathbf{t}^{\mathrm{H}}. \tag{139}$$

By Lemma 2 $f_{\mathcal{R}_C}(\mathbf{v}\mathbf{v}^{\mathrm{H}}) \neq \mathbf{0}$ if and only if $\mathbf{v} = \mathbf{0}$. Since $\mathbf{v} \overset{\mathrm{a.s}}{=} 0$, we have $\gamma \overset{\mathrm{a.s}}{>} 0$, which gives the result. ■

C. PROOF THAT 74 DO NOT GENERATE $-I_{2^{2K+1}}$

The proof that $-\mathbf{I}_{2^{2k+1}}$ is not generated by the operators in (74) is quite tedious. Recall that the stabilizer group generators S^k defined in (74) have $S^k \subset \mathcal{P}_{2k+1}$ and that S^k contains $|S^k| = 2k$ elements. Let G^k denote the group generated by the set of matrices $S^k = \{\mathbf{S}_0^k, \ldots, \mathbf{S}_{2k-1}^k\}$ defined in (74).

Proof: We will proceed by induction on k. The base case, when $k = 1$, can be shown by brute force. For all $\mathbf{G} \in G^1$ for some q, r we can write

$$\begin{aligned}\mathbf{G} &= \mathbf{X}^{q+r} \otimes \mathbf{X}^q\mathbf{Z}^r \otimes \mathbf{X}^q\mathbf{Z}^r \\ &= \mathbf{X}^{q+r \mod 2} \otimes \mathbf{X}^{q \mod 2}\mathbf{Z}^{r \mod 2} \otimes \mathbf{X}^{q \mod 2}\mathbf{Z}^{r \mod 2}.\end{aligned} \tag{140}$$

This follows from the commutativity and the observation that even powers of $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are identity. We will assume that G^p does not contain $-\mathbf{I}_{2^{2p+1}}$, and show that this implies $-\mathbf{I}_{2^{2(p+1)+1}} \notin G^{p+1}$.

Recall the recursive structure of the stabilizer group generators (75). Additionally, recall that a fundamental property of Kronecker products is that $\exists$ a permutation matrix, $\mathbf{D} \in \mathbb{R}^{qr \times qr}$ such that for all $\mathbf{A} \in \mathbb{C}^{q \times q}$, $\mathbf{B} \in \mathbb{C}^{r \times r}$, we have $\mathbf{A} \otimes \mathbf{B} = \mathbf{D}(\mathbf{B} \otimes \mathbf{A})\mathbf{D}^{\mathrm{T}}$. Thus, by (75), for all $i \geq 2$ there exists a permutation matrix $\mathbf{D} \in \mathbb{R}^{2^{2(p+1)+1}}$ we have,

$$\mathbf{S}_i^{p+1} = \mathbf{D}\big(\mathbf{I}_4 \otimes \mathbf{S}_{i-2}^p\big)\mathbf{D}^{\mathrm{T}}. \tag{141}$$

We first show that if any product of elements from S^{p+1} generates $-\mathbf{I}_{2^{2(p+1)+1}}$, it must necessarily include $\mathbf{S}_0^{p+1}$ and/or $\mathbf{S}_1^{p+1}$. We proceed by contradiction. Assume that there exist integers $z_2, z_3, \ldots, z_{2(p+1)-1}$ such that

$$\prod_{i=2}^{2(p+1)-1} \left(\mathbf{S}_i^{p+1}\right)^{z_i} = -\mathbf{I}_{2^{2(p+1)+1}}. \tag{142}$$

Substituting (141) into (142) and recalling that, since $\mathbf{D}$ is a permutation matrix, $\mathbf{D}\mathbf{D}^{\mathrm{T}} = \mathbf{I}_{2^{2(k+1)+1}}$, implies that

$$\begin{aligned}\mathbf{I}_4 \otimes \prod_{i=2}^{2(p+1)-1} \left(\mathbf{S}_{i-2}^p\right)^{z_i} &= -\mathbf{I}_4 \otimes \mathbf{I}_{2^{2p+1}} \\ &= -\mathbf{I}_{2^{2(p+1)+1}},\end{aligned}$$

e.g., that a product of matrices in S^p can generate $-\mathbf{I}_{2^{2p+1}}$. This is impossible by the inductive assumption that $-\mathbf{I}_{2^{2p+1}} \notin G^p$.

Thus, any product of elements in S^{p+1} that is equal to $-\mathbf{I}_{2^{2(p+1)+1}}$ must include a nonzero power of $\mathbf{S}_0^{p+1}$ or $\mathbf{S}_1^{p+1}$. Since for z even $(\mathbf{S}_0^{p+1})^z = (\mathbf{S}_1^{p+1})^z = \mathbf{I}_{2^{2(p+1)+1}}$, the product must contain at least one odd power $\mathbf{S}_0^{p+1}$ or $\mathbf{S}_1^{p+1}$. We have by definition (cf. (74))

$$\mathbf{S}_0^{p+1} = \mathbf{X} \otimes \mathbf{X} \otimes \mathbf{I}_{2^p} \otimes \mathbf{X} \otimes \mathbf{I}_{2^p} \tag{143a}$$
$$\mathbf{S}_1^{p+1} = \mathbf{X} \otimes \mathbf{Z} \otimes \mathbf{I}_{2^p} \otimes \mathbf{Z} \otimes \mathbf{I}_{2^p}. \tag{143b}$$

One can show that $(\mathbf{S}_0^{p+1})^{z_0}(\mathbf{S}_1^{p+1})^{z_1} \neq -\mathbf{I}_{2^{2(p+1)+1}}$ for any (integer) choice of z_0 and z_1 directly via the base case.

Thus, if there exists $z_0, z_1 \ldots z_{2(p+1)-1}$ such that

$$\prod_{i=0}^{2(p+1)-1} \left(\mathbf{S}_i^{p+1}\right)^{z_i} = -\mathbf{I}_{2^{2(p+1)+1}}, \tag{144}$$

there must exist at least one odd z_i for $i \leq 1$ and at least one odd z_i for $i > 1$. In other words, any product generating $-\mathbf{I}_{2^{2(p+1)+1}}$ must include at least one odd power of $\{\mathbf{S}_0^{p+1}, \mathbf{S}_1^{p+1}\}$ and at least one odd power of $\{\mathbf{S}_2^{p+1}, \ldots \mathbf{S}_{2(p+1)-1}^{p+1}\}$.

We will again proceed by contradiction. Assume there exists integers $z_0, z_1 \ldots z_{2(p+1)-1}$ such that (144) holds. Let

$$\mathbf{A} = \left(\mathbf{S}_0^{p+1}\right)^{z_0}\left(\mathbf{S}_1^{p+1}\right)^{z_1}, \tag{145}$$

and let

$$\mathbf{B} = \prod_{i=2}^{2(p+1)-1} \left(\mathbf{S}_i^{p+1}\right)^{z_i}. \tag{146}$$

Let

$$r = \sum_{i=2}^{2(p+1)-1} z_i. \tag{147}$$

The structure of the group generators in (74) and the fact that for $i \leq 1$ at least one of the z_i is odd guarantees that for some $\mathbf{R} \in \mathcal{P}_1$ with $\mathbf{R} \neq \mathbf{I}_2$

$$\mathbf{A} = \mathbf{X}^{(z_0+z_1 \mod 2)} \otimes \mathbf{R} \otimes \mathbf{I}_{2^p} \otimes \mathbf{R} \otimes \mathbf{I}_{2^p}. \tag{148}$$

Analogously, we must have for some $\mathbf{W} \in \mathcal{P}_p$ with $\mathbf{W} \neq \mathbf{I}_{2^p}$

$$\mathbf{B} = \mathbf{X}^{(r \mod 2)} \otimes \mathbf{I}_2 \otimes \mathbf{W} \otimes \mathbf{I}_2 \otimes \mathbf{W}. \tag{149}$$

It turns out that this is all we need for a contradiction. Assuming (144) implies $\mathbf{AB} = -\mathbf{I}_{2^{2(p+1)+1}}$ and thus $\mathbf{B} = -\mathbf{A}^{-1}$. However, since the matrices $S^{p+1} \subset \mathcal{P}_{2(p+1)+1}$ are Hermitian and unitary, we have $\mathbf{A}^{-1} = \mathbf{A}$. Thus (144) implies $\mathbf{B} = -\mathbf{A}$. Since $\mathbf{B}$ is also Hermetian and unitary, $\mathbf{B} = \mathbf{B}^{-1}$ and thus by definition

$$\mathrm{Tr}(\mathbf{BB}) = 2^{2(p+1)+1}. \tag{150}$$

However, setting $d = r + z_0 + z_1$

$$-\mathbf{AB} = \mathbf{X}^{(d \mod 2)} \otimes \mathbf{R} \otimes \mathbf{W} \otimes \mathbf{R} \otimes \mathbf{W}. \tag{151}$$

If $\mathbf{R} \in \mathcal{P}_1$ with $\mathbf{R} \neq \mathbf{I}_2$, then $\mathrm{Tr}(\mathbf{R}) = 0$ (cf. (5)). By the trace property of the Kroneker product

$$\begin{aligned}\mathrm{Tr}(-\mathbf{AB}) &= -\mathrm{Tr}\Big(\mathbf{X}^{(d \mod 2)}\Big)(\mathrm{Tr}(\mathbf{R}))^2(\mathrm{Tr}(\mathbf{W}))^2 &(152)\\ &= 0 &(153)\\ &\neq 2^{2(p+1)+1}, &(154)\end{aligned}$$

which contradicts (150), and thus completes the proof. ■

D. PROOF OF THE SYNDROME PROPERTY

In this Appendix we prove that the stabilizer group generators in S^k (constructed in (74)) have a unique set of commutation relationships with respect to the errors in E^k (defined in (73)). Furthermore, we will demonstrate that the syndrome of $\mathbf{E}_i \in E^k$ with respect to the stabilizer group generators S^k is exactly the big-endian binary expansion of i. As in Appendix C, we will proceed by recursion on k.

We first formally define the property we are trying to prove, beginning by recalling some definitions and properties of the Pauli group. The commutator between two operators $\mathbf{A}$ and $\mathbf{B}$ is denoted $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. The commutator $[\mathbf{A}, \mathbf{B}]$ is zero if $\mathbf{A}$ and $\mathbf{B}$ commute. If $\mathbf{A}$ and $\mathbf{B}$ are both Pauli operators, then $[\mathbf{A}, \mathbf{B}] \neq \mathbf{0}$ implies that $[\mathbf{A}, \mathbf{B}] = 2\mathbf{AB}$ since such operators either commute or anti-commute. By definition, $E^k \subset \mathcal{P}_{2k+1}$ for all k. To clarify the notation, for fixed k, denote the i^{th} bit of the syndrome of error q with respect to the stabilizer group generators S^k as (cf. (17))

$$s_i^k(q) = \left[s_{C_{S^k}}\left(\mathbf{E}_q^k\right)\right]_i, \tag{155}$$

where $q \in \{0, 1, \ldots, 4^k-1\}$ and $i \in \{0, 1, \ldots, 2k-1\}$. Using the definition of the commutator, we have (cf. (17))

$$s_i^k(q) = \begin{cases} 0 & \left[\mathbf{S}_i^k, \mathbf{E}_q^k\right] = \mathbf{0}_{2^{2k+1}} \\ 1 & \text{otherwise,} \end{cases} \tag{156}$$

i.e., the bit is 0 if $\mathbf{E}_q$ commutes with $\mathbf{S}_i$ and is 1 if the aforementioned operators anticommute. We will prove that, for all k,

$$s_i^k(q) \text{ is the } i^{\text{th}} \text{ bit in the big-endian binary expansion of } q. \tag{157}$$

Per Section V-A, this property holds when $k = 1$. We assume that this property hold for $k = p$, and demonstrate that this implies it holds for $k = p + 1$.

Again, we invoke the recursive property of the stabilizer group generators in (75), as well as a recursive definition of the error operators. Consider equations (71) and (72). These equations, coupled with the properties of the 4-ary expansions of natural numbers give that for $q \in \{0, 1, \ldots, 4^{p+1} - 1\}$:

$$\mathbf{\Gamma}_{p+1}(q) = \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \otimes \mathbf{\Gamma}_p\left(q \bmod 4^p\right), \tag{158}$$

and thus we have by (73) that

$$\mathbf{E}_q^{p+1} = \mathbf{I}_{2^{p+2}} \otimes \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \otimes \mathbf{\Gamma}_p\left(q \bmod 4^p\right). \tag{159}$$

It helps to expand $\mathbf{E}_q^{p+1}$ via

$$\mathbf{E}_q^{p+1} = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{I}_{2^p} \otimes \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \otimes \mathbf{\Gamma}_p\left(q \bmod 4^p\right). \tag{160}$$

Recall that for $i > 1$, for some $\mathbf{W}_i \in \mathcal{P}_p$, (cf. (75))

$$\mathbf{S}_{i-2}^p = \mathbf{X} \otimes \mathbf{W}_i \otimes \mathbf{W}_i \tag{161a}$$

$$\mathbf{S}_i^{p+1} = \mathbf{X} \otimes \mathbf{I}_2 \otimes \mathbf{W}_i \otimes \mathbf{I}_2 \otimes \mathbf{W}_i. \tag{161b}$$

Define $\mathbf{F} = \mathbf{X} \otimes \mathbf{I}_2$. For $i > 1$, by (160) and (161)

$$\mathbf{S}_i^{p+1}\mathbf{E}_q^{p+1} = \mathbf{F} \otimes \mathbf{W}_i \otimes \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \otimes \left(\mathbf{W}_i\mathbf{\Gamma}_p\left(q \bmod 4^p\right)\right). \tag{162}$$

Using an analogous expression for $\mathbf{E}_q^{p+1}\mathbf{S}_i^{p+1}$ and invoking the bilinearity of the Kronecker product, we have the following expression for the commutator

$$\begin{aligned}&\left[\mathbf{S}_i^{p+1}, \mathbf{E}_q^{p+1}\right] \\ &= \mathbf{F} \otimes \mathbf{W}_i \otimes \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \otimes \left[\mathbf{W}_i, \mathbf{\Gamma}_p\left(q \bmod 4^p\right)\right]\end{aligned} \tag{163}$$

where we have "moved the commutator operator inside the Kronecker product." Note that since $\mathbf{F} \in \mathcal{P}_2$, $\mathbf{W}_i \in \mathcal{P}_n$, and $\mathbf{\Omega}(\lfloor \frac{q}{4^p} \rfloor) \in \mathcal{P}_1$, we have

$$\mathbf{F} \otimes \mathbf{W}_i \otimes \mathbf{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right) \neq \mathbf{0}_{2^{p+3}}. \tag{164}$$

As a consequence, it must be that

$$\begin{aligned}\left[\mathbf{S}_i^{p+1}, \mathbf{E}_q^{p+1}\right] &= \mathbf{0}_{2^{2(p+1)+1}} \text{ if and only if} \\ \left[\mathbf{W}_i, \mathbf{\Gamma}_p\left(q \bmod 4^p\right)\right] &= \mathbf{0}_{2^p}.\end{aligned} \tag{165}$$

Thus, for $i > 1$, we can adapt (156) to the case of $k = p + 1$ via

$$s_i^{p+1}(q) = \begin{cases} 0 & \left[\mathbf{W}_i, \mathbf{\Gamma}_p(q \bmod 4^p)\right] = \mathbf{0}_{2^p} \\ 1 & \text{otherwise.} \end{cases} \tag{166}$$

We will relate this expression to the syndrome of the errors in E^p with respect to the stabilizer group generators S^p.

First, note that for all $z \in \{0, 1, \ldots, 4^p - 1\}$, we have

$$\mathbf{E}_z^p = \mathbf{I}_2 \otimes \mathbf{I}_{2^p} \otimes \mathbf{\Gamma}_p(z). \tag{167}$$

Let $z = q \bmod 4^p$. Note that by definition $z \in \{0, 1, \ldots, 4^p - 1\}$. Again assuming $i > 1$ and recalling (161), the product $\mathbf{S}_{i-2}^p$ and $\mathbf{E}_z^p$ is given by

$$\mathbf{S}_{i-2}^p\mathbf{E}_z^p = \mathbf{X} \otimes \mathbf{W}_i \otimes \left(\mathbf{W}_i\mathbf{\Gamma}_p(z)\right) \tag{168}$$

Again, a similar expression for $\mathbf{E}_z^p\mathbf{S}_{i-2}^p$ and the bilinearity of the Kronecker product gives an expression for the commutator

$$\left[\mathbf{S}_{i-2}^p, \mathbf{E}_z^p\right] = \mathbf{X} \otimes \mathbf{W}_i \otimes \left(\left[\mathbf{W}_i, \mathbf{\Gamma}_p(z)\right]\right). \tag{169}$$

In direct analogy to (166), we have (again, for $i > 1$)

$$s_{i-2}^p\left(q \bmod 4^p\right) = \begin{cases} 0, & \left[\mathbf{W}_i, \mathbf{\Gamma}_p(q \bmod 4^p)\right] = \mathbf{0}_{2^p} \\ 1, & \text{otherwise,} \end{cases} \tag{170}$$

where we substituted in the value of z. Comparing (170) with (166), the stabilizer generators recurrence relationship (161) implies

$$s_{i-2}^p\left(q \bmod 4^p\right) = s_i^{p+1}(q). \tag{171}$$

By our inductive assumption, $s_{i-2}^p(q \bmod 4^p)$ is exactly the $(i - 1)^{\text{th}}$ bit in the binary expansion of $q \bmod 4^p$. This is enough to conclude the following claim:

Claim 1: For all $i > 1$, $s_i^{p+1}(q)$ is the i^{th} bit in the big-endian binary expansion of q.

Proof of Claim 1: The claim can be gleaned from some manipulations of big-endian binary expansions of nonnegative integers. Let $q \in \{0, 1, \ldots, 4^{p+1} - 1\}$. In particular for $i > 1$, the i^{th} bit of the $2(p + 1)$ bit expansion of q is the same as the i^{th} bit of the $2(p+1)$ bit expansion of $q \bmod 4^p$. Formally, if q has the $2(p + 1)$ bit expansion

$$q = \left(q_0q_1q_2 \cdots q_{2(p+1)-1}\right)_2, \tag{172}$$

where q_0 is the most significant bit, then $q \bmod 4^p$ has the $2(p + 1)$ bit binary expansion

$$q \bmod 4^p = \left(00q_2q_3 \cdots q_{2(p+1)-1}\right)_2. \tag{173}$$

The $2p$ bit binary expansion of $q \bmod 4^p$ is the same as (173), but with the leading zeros truncated, e.g.,

$$q \bmod 4^p = \left(q_2q_3 \cdots q_{2(p+1)-1}\right)_2. \tag{174}$$

TABLE 2. The commutation relations between stabilizer group generators S_0^{p+1} and S_1^{p+1} and error operators E_q^{p+1} can be written as a function of $\lfloor \frac{q}{4^p} \rfloor$. These follow immediately from (178) and the definition of the matrix valued function Ω in (71).

Commutation Relationships		
	$s_0^{p+1}(q)$	$s_1^{p+1}(q)$
$\lfloor \frac{q}{4^p} \rfloor = 0$	0	0
$\lfloor \frac{q}{4^p} \rfloor = 1$	0	1
$\lfloor \frac{q}{4^p} \rfloor = 2$	1	0
$\lfloor \frac{q}{4^p} \rfloor = 3$	1	1

In other words, if $i > 1$, the i^{th} bit in the $2(p+1)$ bit big-endian expansion of q is the same as the $(i-2)^{\text{th}}$ bit in the $2p$ bit big-endian expansion of q mod 4^p. The inductive assumption implies that

$$s_{i-2}^{p}\left(q \bmod 4^p\right) = \left(q_2 q_3 \dots q_{2(p+1)-1}\right)_2. \tag{175}$$

Thus by (171), the claim holds. ■

Claim 1 demonstrates that for $i > 2$, $s_i^{p+1}(q)$ is the i^{th} bit in the binary expansion of q. Agreement on the first two bits essentially follows from the base case. Note that if q has the expansion in (172),

$$\left\lfloor \frac{q}{4^p} \right\rfloor = (q_0 q_1)_2. \tag{176}$$

Recall the forms of $\mathbf{S}_0^{p+1}$ and $\mathbf{S}_1^{p+1}$ in (143). Let $\mathbf{J} = \boldsymbol{\Gamma}_p(q \bmod 4^p)$. Using (160), we can write

$$\left[\mathbf{S}_0^{p+1}, \mathbf{E}_q^{p+1}\right] = \mathbf{X} \otimes \mathbf{X} \otimes \mathbf{I}_{2^p} \otimes \left[\mathbf{X}, \boldsymbol{\Omega}\left(\left\lfloor \frac{q}{4^p} \right\rfloor\right)\right] \otimes \mathbf{J} \tag{177a}$$

$$\left[\mathbf{S}_1^{p+1}, \mathbf{E}_q^{p+1}\right] = \mathbf{X} \otimes \mathbf{Z} \otimes \mathbf{I}_{2^p} \otimes \left[\mathbf{Z}, \boldsymbol{\Omega}\left(\left\lfloor \frac{q}{4^p} \right\rfloor\right)\right] \otimes \mathbf{J}. \tag{177b}$$

Thus, in analog to (166) and (170) we have

$$s_0^{p+1}(q) = \begin{cases} 0 & \left[\mathbf{X}, \boldsymbol{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right)\right] = \mathbf{0}_2 \\ 1 & \text{otherwise.} \end{cases} \tag{178a}$$

$$s_1^{p+1}(q) = \begin{cases} 0 & \left[\mathbf{Z}, \boldsymbol{\Omega}\left(\lfloor \frac{q}{4^p} \rfloor\right)\right] = \mathbf{0}_2 \\ 1 & \text{otherwise.} \end{cases} \tag{178b}$$

We need to show that $s_0^{p+1}(q) = q_0$ and $s_1^{p+1}(q) = q_1$. Table 2 gives $s_0^{p+1}(q)$ and $s_1^{p+1}(q)$ as a function of $\lfloor \frac{q}{4^p} \rfloor$. The binary expansion of $\lfloor \frac{q}{4^p} \rfloor$ is exactly

$$\left\lfloor \frac{q}{4^p} \right\rfloor = \left(s_0^{p+1}(q) s_1^{p+1}(q)\right)_2. \tag{179}$$

Comparing with (176) shows agreement on the first two bits and completes the proof. ■

E. CHERNOFF BOUND

Our approach to deriving the bound on the pairwise error probability mirrors that in [18]. We assume the decoding rule in (111). Define $\mathbf{q}_i$ and $\mathbf{w}_i$ as in (101). Let $\hat{\mathbf{w}}_i = [\hat{c}_i, \ \frac{\sqrt{2}}{\sigma}\mathbf{w}_i^{\mathrm{T}}]^{\mathrm{T}}$. The assumed and derived independence relationships in Section VI guarantee that $\hat{\mathbf{w}}_i \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_3)$ IID. Note that by definition, given the random transmitted symbol $\mathbf{s}_{\mathrm{TX}} \in \mathcal{C}$, we have (cf. (101))

$$\mathbf{q}_i = \left[\mathbf{s}_{\mathrm{TX}} \ \frac{\sigma}{\sqrt{2}}\mathbf{I}_2\right]\hat{\mathbf{w}}_i. \tag{180}$$

We first derive the *pairwise error probability*. Let

$$\gamma_i = \mathbf{s}_p^{\mathrm{H}}\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}\mathbf{s}_p - \mathbf{s}_q^{\mathrm{H}}\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}\mathbf{s}_q \tag{181}$$

and define

$$\gamma = \sum_{i=0}^{4^k-1} \gamma_i. \tag{182}$$

The error event of detecting $\mathbf{s}_q$ when the transmitted symbol was $\mathbf{s}_p$, assuming only (those) two hypotheses, is given by (cf. (111))

$$\mathbb{P}\left[\mathbf{s}_p \rightarrow \mathbf{s}_q\right] = \mathbb{P}\left[\gamma \leq 0 | \mathbf{s}_{\mathrm{TX}} = \mathbf{s}_p\right]. \tag{183}$$

Note that the γ_i are mutually independent and identically distributed. This error event is amenable to analysis via the Chernoff bound, which requires computing the moment generating function of γ_i [18], [48].

Assume $\mathbf{s}_{\mathrm{TX}} = \mathbf{s}_p$. Let $\mathbf{r}_i = [r_{i,1}, \ r_{i,2}]^{\mathrm{T}} \in \mathbb{C}^2$ be given by

$$\mathbf{r}_i = \begin{bmatrix} \mathbf{s}_p^{\mathrm{H}} \\ \mathbf{s}_q^{\mathrm{H}} \end{bmatrix}\mathbf{q}_i. \tag{184}$$

Writing $\mathbf{q}_i$ in terms of $\hat{\mathbf{w}}_i$ via (180) gives

$$\mathbf{r}_i = \begin{bmatrix} 1 & \frac{\sigma}{\sqrt{2}}\mathbf{s}_p^{\mathrm{H}} \\ \mathbf{s}_q^{\mathrm{H}}\mathbf{s}_p & \frac{\sigma}{\sqrt{2}}\mathbf{s}_q^{\mathrm{H}} \end{bmatrix}\hat{\mathbf{w}}_i. \tag{185}$$

Let $\delta_{p,q} = \mathbf{s}_p^{\mathrm{H}}\mathbf{s}_q$. Letting

$$\mathbf{K} = \begin{bmatrix} 1 + \frac{\sigma^2}{2} & \delta_{p,q}\left(1 + \frac{\sigma^2}{2}\right) \\ \delta_{p,q}^{\mathrm{H}}\left(1 + \frac{\sigma^2}{2}\right) & |\delta_{p,q}|^2 + \frac{\sigma^2}{2} \end{bmatrix}, \tag{186}$$

we have that the $\mathbf{r}_i$ are IID with $\mathbf{r}_i \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{K})$. Recall that

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{187}$$

We can express γ_i via

$$\gamma_i = \mathbf{r}_i^{\mathrm{H}}\mathbf{Z}\mathbf{r}_i, \tag{188}$$

which immediately follows from (181) and (184).

We compute the relevant moment generating function via

$$\Gamma(t) = \mathbb{E}_{\mathbf{r}}\left[e^{-t\gamma_i}\right] = \iint_{\mathbb{C}^2} \frac{e^{-t\mathbf{r}^{\mathrm{H}}\mathbf{Z}\mathbf{r} - \mathbf{r}^{\mathrm{H}}\mathbf{K}^{-1}\mathbf{r}}}{\det(\mathbf{K})\pi^2} dr_1 dr_2 \tag{189}$$

$$= \iint_{\mathbb{C}^2} \frac{e^{-\mathbf{r}^{\mathrm{H}}(t\mathbf{Z} + \mathbf{K}^{-1})\mathbf{r}}}{\det(\mathbf{K})\pi^2} dr_1 dr_2, \tag{190}$$

where $\mathbf{r} = [r_1, r_2]^{\mathrm{T}}$. Let $\Re(t) = t_r$, and define the matrix $\boldsymbol{\Upsilon}$ as

$$\boldsymbol{\Upsilon} = \mathbf{K}^{-1} + t_r\mathbf{Z}. \tag{191}$$

For the integral (189) to converge, we require that

$$\boldsymbol{\Upsilon} \succ \mathbf{0}_2, \tag{192}$$

which ensures that the real part of the argument to the exponential is negative as $|r_1|, |r_2| \to \infty$. Let

$$v_- = \frac{\left(1 - |\delta_{p,q}|^2\right) - \sqrt{\left(1 - |\delta_{p,q}|^2\right)^2 + 4 \det \mathbf{K}}}{2 \det \mathbf{K}} \tag{193}$$

and

$$v_+ = \frac{\left(1 - |\delta_{p,q}|^2\right) + \sqrt{\left(1 - |\delta_{p,q}|^2\right)^2 + 4 \det \mathbf{K}}}{2 \det \mathbf{K}}. \tag{194}$$

Define the radius of convergence

$$\mathrm{ROC} = (v_-, v_+). \tag{195}$$

With some work, it can be shown that we have (192) so long as $t_r \in \mathrm{ROC}$. When $t_r \in \mathrm{ROC}$ we have

$$\Gamma(t) = \frac{1}{\det(\mathbf{K}) \det\left(\mathbf{K}^{-1} + t\mathbf{Z}\right)}. \tag{196}$$

From the Chernoff bound, since (182) is the sum of 4^k IID random variables, we have (cf. [18], [48])

$$\mathbb{P}\left[\mathbf{s}_p \to \mathbf{s}_q\right] \leq \min_{t \in \mathbb{R}^+} (\Gamma(t))^{4^k}. \tag{197}$$

It is clear that the t which minimizes (197) is the t that maximizes $\det(\mathbf{K}^{-1} + t\mathbf{Z})$. The maximizer is given by

$$\hat{t} = \frac{1}{2\frac{\sigma^2}{2}\left(1 + \frac{\sigma^2}{2}\right)}, \tag{198}$$

which lies within the radius of convergence. The achieved minimum is

$$\Gamma(\hat{t}) = \frac{4\left(\frac{\sigma^2}{2}\right)\left(1 + \frac{\sigma^2}{2}\right)}{4\left(\frac{\sigma^2}{2}\right)\left(1 + \frac{\sigma^2}{2}\right) + 1 - |\delta_{p,q}|^2}. \tag{199}$$

We define the signal-to-noise ratio as the ratio of average transmitted power to received noise variance via $\mathrm{SNR} = \frac{1}{\sigma^2}$ and we denote its inverse via $\xi = \sigma^2 = 1/\mathrm{SNR}$. Substituting ξ into (199), and substituting the latter into (197) gives the bound

$$\mathbb{P}\left[\mathbf{s}_p \to \mathbf{s}_q\right] \leq \left(1 - \frac{1 - |\delta_{p,q}|^2}{\xi^2 + 2\xi + \left(1 - |\delta_{p,q}|^2\right)}\right)^{4^k}. \tag{200}$$

As may be expected, (200) is an increasing function of $|\delta_{p,q}|^2 = |\mathbf{s}_p^{\mathrm{H}} \mathbf{s}_q|^2$. Let

$$|\hat{\delta}|^2 = \max_{\substack{\mathbf{s}_q, \mathbf{s}_r \in \mathcal{C} \\ \mathbf{s}_q \neq \mathbf{s}_r}} |\mathbf{s}_q^{\mathrm{H}} \mathbf{s}_r|^2. \tag{201}$$

We have, for all $p \neq q$

$$\mathbb{P}\left[\mathbf{s}_p \to \mathbf{s}_q\right] \leq \left(1 - \frac{1 - |\hat{\delta}|^2}{\xi^2 + 2\xi + \left(1 - |\hat{\delta}|^2\right)}\right)^{4^k}. \tag{202}$$

We can union bound the probability of error P_e via the sum of all pairwise error probabilities via

$$P_e \leq \sum_{\mathbf{s}_p \in \mathcal{C}} \mathbb{P}\left[\mathbf{s}_{\mathrm{TX}} = \mathbf{s}_p\right] \sum_{\substack{\mathbf{s}_q \in \mathcal{C} \\ \mathbf{s}_q \neq \mathbf{s}_p}} \mathbb{P}\left[\mathbf{s}_p \to \mathbf{s}_q\right]. \tag{203}$$

Assuming that the distribution transmitted symbols are IID uniform over $\mathcal{C}$ and applying the bound (202) gives

$$P_e \leq (|C| - 1)\left(1 - \frac{1 - |\hat{\delta}|^2}{\xi^2 + 2\xi + \left(1 - |\hat{\delta}|^2\right)}\right)^{4^k}. \tag{204}$$

The diversity of the space-time code can be deduced by computing the first order Taylor expansion of (202) at the infinite SNR limit (i.e., around $\xi = 0$). Substituting this into (204), we obtain

$$P_e \lessapprox (|\mathcal{C}| - 1)\left(\frac{2\xi}{1 - |\hat{\delta}|^2}\right)^{4^k}, \text{ at high SNR.} \tag{205}$$

Letting $\mathrm{SNRdB} = 10 \log_{10}(\mathrm{SNR})$, and

$$\beta = \log_{10}(|\mathcal{C}| - 1) + 4^k \log_{10}\left(\frac{2}{1 - |\hat{\delta}|^2}\right), \tag{206}$$

we have

$$\log_{10}(P_e) \lessapprox \beta - \tfrac{4^k}{10}(\mathrm{SNRdB}), \text{ at high SNR,} \tag{207}$$

implying that at high SNR, a 10 dB increase in SNR results in a 4^k decade drop in the probability of error. The $M \times M = 2^k \times 2^k$ channel has exactly 4^k complex degrees of freedom, so it is evident that the space-time code achieves full diversity.

ACKNOWLEDGMENT

The authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions. Travis C. Cuvelier is grateful for additional suggestions from K. U. Mazher.

REFERENCES

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[2] M. M. Wilde, “From classical to quantum Shannon theory,” 2019, *arXiv:1106.1445*.

[3] D. Gottesman, “Stabilizer codes and quantum error correction,” Ph.D. dissertation, Div. Phys. Math. Astron., California Inst. Technol., Pasadena, CA, USA, 1997. [Online]. Available: https://thesis.library.caltech.edu/2900/

[4] A. R. Calderbank, E. M. Rains, P. Shor, and N. J. A. Sloane, “Quantum error correction via codes over GF(4),” *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1369–1387, Jul. 1998.

[5] D. J. MacKay, G. Mitchison, and P. L. McFadden, “Sparse-graph codes for quantum error correction,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2315–2330, Oct. 2004.

[6] Y. C. Eldar and A. V. Oppenheim, “Quantum signal processing,” *IEEE Signal Process. Mag.*, vol. 19, no. 6, pp. 12–32, Nov. 2002.

[7] Y. C. Eldar, “Quantum signal processing,” Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2001. [Online]. Available: https://dspace.mit.edu/bitstream/handle/1721.1/16805/50544999-MIT.pdf?sequence=2&isAllowed=y

[8] B. R. L. Cour and G. E. Ott, "Signal-based classical emulation of a universal quantum computer," *New J. Phys.*, vol. 17, May 2015, Art. no. 053017. [Online]. Available: https://iopscience.iop.org/article/10.1088/1367-2630/17/5/053017

[9] C. Ostrove, B. R. L. Cour, A. Lanham, and G. E. Ott, "Improving performance of an analog electronic device using quantum error correction," *J. Phys. Commun.*, vol. 3, no. 8, 2019, Art. no. 085017. [Online]. Available: https://iopscience.iop.org/article/10.1088/2399-6528/ab3c37

[10] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.

[11] R. W. Heath, Jr. and A. Lozano, *Foundations of MIMO Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2019.

[12] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[13] R. W. Heath and A. J. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 962–968, Jun. 2005.

[14] A. Lozano and N. Jindal, "Transmit diversity vs. spatial multiplexing in modern mimo systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 186–197, Jan. 2010.

[15] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[16] O. N. Yilmaz, Y.-P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, 2015, pp. 1190–1195.

[17] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, Feb. 2015.

[18] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[19] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[20] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000.

[21] D. Agrawal, T. J. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2618–2626, Sep. 2001.

[22] W. Zhao, G. Leus, and G. B. Giannakis, "Orthogonal design of unitary constellations for uncoded and trellis-coded noncoherent space-time systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1319–1327, Jun. 2004.

[23] D. Xia, J.-K. Zhang, S. Dumitrescu, and F.-K. Gong, "Full diversity non-coherent alamouti-based toeplitz space-time block codes," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5241–5253, Oct. 2012.

[24] I. Kammoun and J.-C. Belfiore, "A new family of Grassmann space-time codes for non-coherent MIMO systems," *IEEE Commun. Lett.*, vol. 7, no. 11, pp. 528–530, Nov. 2003.

[25] A. Ashikhmin and A. R. Calderbank, "Grassmannian packings from operator Reed–Muller codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5689–5714, Nov. 2010.

[26] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[27] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1485–1503, Jun. 2002.

[28] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2567–2578, Nov. 2000.

[29] B. L. Hughes, "Optimal space-time constellations from groups," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 401–410, Feb. 2003.

[30] L. H.-J. Lampe, R. Schober, and R. F. H. Fischer, "Coded differential space-time modulation for flat fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 582–590, May 2003.

[31] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2335–2367, Sep. 2001.

[32] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2626–2631, Sep. 2001.

[33] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[34] R. H. Gohary and H. Yanikomeroglu, "Noncoherent MIMO signaling for block-fading channels: Approaches and challenges," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 80–88, Mar. 2019.

[35] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[36] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2399, Sep. 2001.

[37] S. A. Lanham, T. C. Cuvelier, C. Ostrove, B. L. Cour, G. Ott, and R. W. Heath, Jr., "A noncoherent space-time code from quantum error correction," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–6.

[38] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[39] D. Nigg *et al.*, "Quantum computations on a topologically encoded qubit," *Science*, vol. 345, no. 6194, pp. 302–305, 2014. [Online]. Available: https://science.sciencemag.org/content/345/6194/302

[40] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley-Intersci., 2006.

[42] D. Love. "Grassmannian Subspace Packing." [Online]. Available: https://engineering.purdue.edu/∼djlove/grass.html (Accessed: Aug. 1, 2021).

[43] I. S. Dhillon, R. W. Heath, Jr., T. Strohmer, and J. A. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Exp. Math.*, vol. 17, no. 1, pp. 9–35, 2008.

[44] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Exp. Math.*, vol. 5, no. 2, pp. 139–159, 1996.

[45] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[46] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[47] L. C. Tran, T. A. Wysocki, J. Seberry, A. Mertins, and S. S. Adams, "Novel constructions of improved square complex orthogonal designs for eight transmit antennas," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4439–4448, Oct. 2009.

[48] J. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill Companies, 2000.

TRAVIS C. CUVELIER (Student Member, IEEE) received the B.S. and M.Eng. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2015 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Texas at Austin. He previously held internships with LGS Innovations and the MITRE Corporation. At UT, he is affiliated with the Wireless Networking and Communications Group, the Oden Institute for Computational Engineering and Sciences, and the Applied Research Laboratories. His research interests include broad areas of signal processing and information theory with applications to wireless communications, quantum information, and feedback control.

S. ANDREW LANHAM (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from the University of Texas at Austin in 2017 and 2019, respectively, where he has been an Engineering Scientist with Applied Research Laboratories, Center for Quantum Research, since 2019. His research interests are in the broad area of signal processing and information theory, with particular focus on quantum information science.

BRIAN R. LA COUR received the B.S. degree in physics and the M.S. degree in physics and mathematics from the University of New Orleans, New Orleans, LA, USA, in 1991 and 1995, respectively, and the Ph.D. degree in physics from the University of Texas at Austin in 2000, specializing in statistical mechanics.

He is currently a Research Scientist with Applied Research Laboratories, the University of Texas at Austin, where he directs the Center for Quantum Research. His research interests include sonar signal processing and quantum information science.

ROBERT W. HEATH, JR. (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from the University of Virginia, Charlottesville, VA, USA, in 1996 and 1997, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2002.

From 1998 to 2001, he was a Senior Member of the Technical Staff then a Senior Consultant with Iospan Wireless, Inc, San Jose, CA, USA, where he worked on the design and implementation of the physical and link layers of the first commercial MIMO-OFDM communication system. From 2002 to 2020, he was with The University of Texas at Austin, most recently as the Cockrell Family Regents Chair of Engineering and the Director of UT SAVES. He is currently the Lampe Distinguished Professor with North Carolina State University and a Co-Founder of 6GNC. He is also the President and a CEO of MIMO Wireless, Inc. He has authored *Introduction to Wireless Digital Communication* (Prentice Hall, 2017) and *Digital Wireless Communication: Physical Layer Exploration Lab Using the NI USRP* (National Technology and Science Press, 2012), and coauthored *Millimeter Wave Wireless Communications* (Prentice Hall, 2014) and *Foundations of MIMO Communication* (Cambridge University Press, 2018). He has been a coauthor of a number award winning conference and journal papers, including recently the 2017 Marconi Prize Paper Award, the 2019 IEEE Communications Society Stephen O. Rice Prize, the 2020 IEEE Signal Processing Society Overview Paper Award and the 2021 IEEE Vehicular Technology Society Neal Shepherd Memorial Best Propagation Paper Award. He received the 2017 EURASIP Technical Achievement Award and the 2019 IEEE Kiyo Tomiyasu Award and the 2021 IEEE Vehicular Technology Society James Evans Avant Garde Award. He is a Member-at-Large on the IEEE Communications Society Board-of-Governors (2020–2022) and was a past Member-at-Large on the IEEE Signal Processing Society Board-of-Governors (2016–2018). He was the Editor-in-Chief of IEEE Signal Processing Magazine from 2018 to 2020. In 2017, he was selected as a Fellow of the National Academy of Inventors. He is also a licensed Amateur Radio Operator, a Private Pilot, a registered Professional Engineer in Texas.